Department of Psychology and Logopedics

Faculty of Medicine

University of Helsinki

EFFECTS OF VOCAL MUSIC ON VERBAL LEARNING AND LONG-TERM RECOVERY AFTER STROKE

Vera Leo

Doctoral Programme of Psychology,

Learning and Communication (PsyCo)

DOCTORAL DISSERTATION

To be presented for public discussion with the permission of the Faculty of Medicine of the University of Helsinki in Lecture Hall 2, Haartman Institute, Haartmaninkatu 3, on the 2nd of October, 2020 at 12 o'clock

Helsinki 2020

Supervisors Associate Professor Teppo Särkämö, PhD
Department of Psychology and Logopedics
Faculty of Medicine
University of Helsinki

Research Director Mari Tervaniemi, PhD
Cicero Learning Network, Faculty of Educational Sciences *and* Cognitive Brain Research Unit, Department Psychology and Logopedics,
Faculty of Medicine
University of Helsinki

Professor Seppo Soinila, PhD
Department of Clinical Medicine
Faculty of Medicine
University of Turku

Reviewers Professor Sonja Kotz, PhD
Department of Neuropsychology and Psychopharmacology
Faculty of Psychology and Neuroscience
University of Maastricht, NL

Professor Heikki Hämäläinen, PhD
Department of Psychology and Speech-Language Pathology
Faculty of Social Sciences
University of Turku

Opponent Docent Mervi Jehkonen, PhD
Department of Psychology
Faculty of Social Sciences
University of Tampere

The Faculty of Medicine uses the Urkund system (plagiarism recognition) to examine all doctoral dissertations.

ISBN 978-951-51-6447-6 (nid.)
ISBN 978-951-51-6448-3 (PDF)

Unigrafia

Helsinki 2020

Abstract

The prevalence of stroke increases in the ageing population entailing an enormous economic and societal burden. This has raised the need for motivating, effective and easily applicable rehabilitation tools to enhance recovery and neuroplasticity. Music is an important source of enjoyment and well-being across life and it provides a multidomain stimulus that is both pleasant and rewarding, and engages the brain extensively. Previous evidence suggests that daily music listening can enhance cognitive recovery and mood and induce functional and structural neuroplasticity changes after stroke. Songs may also function as a verbal learning aid in healthy subjects. The aim of this thesis was to further explore the specific role of vocal (sung) music as a tool to aid verbal learning and long-term recovery after stroke.

In Studies I and II, stroke patients (N = 31) performed a verbal learning task where novel narrative stories were presented in both spoken and sung formats, and underwent MRI at acute and 6-month post-stroke stages. Study I showed that stroke patients, especially those with mild aphasia, learned and recalled the stories better when they were presented in sung than spoken format at the 6-month stage. Exploring the cognitive and neural mechanisms underlying this effect, Study II further showed that non-aphasic patients exhibited more stable recall, indicated by reduced serial position effects, whereas aphasic patients showed a larger recency effect and enhanced chunking in the sung than spoken task. Diffusion tensor imaging and voxel-based morphometry results indicated that these effects were coupled with greater volume of the left arcuate fasciculus in non-aphasics, and with greater volume of the right inferior fronto-occipital fasciculus and grey matter in a bilateral network of temporal, frontal, and parietal regions in aphasics.

In Study III, data was pooled from two randomized controlled trials where stroke patients (N = 83) received an intervention involving daily listening to self-selected vocal music, instrumental music, or audiobooks during the first three months after stroke. The recovery was assessed with neuropsychological tests and a mood questionnaire at acute, 3-month and 6-month stages, and structural MRI and functional MRI (fMRI) at acute and 6-month stages. Compared to audiobooks, listening to music enhanced the recovery of language skills and verbal memory and reduced negative mood. Vocal music had the strongest rehabilitative effect on both language and verbal memory, and the positive effects of music listening on language recovery were seen especially in patients with aphasia. Results from voxel-based morphometry and resting-state, and task-based fMRI analyses showed that vocal music listening selectively increased grey matter volume in left temporal areas and functional connectivity in the default mode network from acute to 6-month stage.

The findings of the present thesis provide further evidence that listening to vocal music is a useful tool to support cognitive and emotional recovery after stroke and to enhance early language recovery in aphasia. The rehabilitative effects are driven by both structural and functional plasticity changes in temporoparietal networks, which are crucial for emotional processing, language and memory.

Tiivistelmä

Väestön ikääntyessä yhä useampi sairastuu aivoverenkiertohäiriöön (AVH), minkä aiheuttama yksilöllinen ja yhteiskunnallinen haitta on valtava. Tästä johtuen tarvitaan motivoivia, tehokkaita ja helposti saatavilla olevia työkaluja tehostamaan kuntoutusta ja edesauttamaan aivojen muovautuvuutta toipumisvaiheessa. Musiikki on tärkeä nautinnon ja hyvinvoinnin lähde ja monipuolinen virike, joka miellyttää, palkitsee ja aktivoi aivoja laajalti. Aiemmissa tutkimuksissa on havaittu, että päivittäinen musiikin kuuntelu AVH:n jälkeisten kuukausien aikana tehostaa kognitiivisten toimintojen ja mielialan kuntoutumista ja saa aikaan toiminnallista ja rakenteellista muovautuvuutta otsa- ja ohimolohkoalueilla sekä limbisillä alueilla, ja että laulut toimivat kielellisen oppimisen tukena terveillä henkilöillä. Tässä väitöskirjassa tarkastellaan erityisesti lauletun musiikin vaikutusta kielelliseen oppimiseen sekä pitkäkestoiseen toipumiseen AVH:n jälkeen.

Tutkimuksissa I ja II AVH-potilaille (N = 31) tehtiin kielellinen oppimistehtävä, jossa heille esitettiin uusia tarinoita sekä laulettuna että puhuttuna, ja aivojen rakenteellinen magneettikuvaus (MRI) akuuttivaiheessa ja 6 kk sairastumisen jälkeen. Tutkimus I osoitti, että erityisesti ne potilaat, joilla oli lievä afasia, oppivat ja muistivat toipumisvaiheessa 6 kk sairastumisen jälkeen laulettuna esitetyn tarinan paremmin verrattuna puhuttuna esitettyyn. Tutkimus II selvitti tämän taustalla olevia kognitiivisia ja neuraalisia mekanismeja ja osoitti, että ei-afaattiset potilaat muistivat lauletun tarinan puhuttua tasaisemmin, mikä näkyi pienentyneenä sarjapositiovaikutuksena, kun taas afasiapotilailla lauletun tarinan muistamisessa ilmeni suurempi äskeisyysvaikutus ja tehokkaampi mieltämisyksiköiden muodostaminen (engl. chunking). Diffuusiotensorikuvantamisella ja vokselipohjaisella morfometrialla (VBM) saadut tulokset osoittivat, että nämä efektit olivat yhteydessä vasemman arcuate fasciculus (AF) –radaston tilavuuteen ei-afasiapotilailla ja oikean inferior fronto-occipital fasciculus (IFOF) –radaston sekä bilateraalisten ohimo-, otsa- ja päälakilohkoalueiden tilavuuteen afasiapotilailla.

Tutkimuksessa III yhdistettiin kahden satunnaistetun kontrolloidun tutkimuksen AVH-potilaiden aineistot (N = 83) ja tutkittiin, miten kahden kuukauden ajan tapahtuva päivittäinen laulumusiikin, instrumentaalimusiikin tai äänikirjojen kuuntelu vaikuttaa toipumiseen. Toipumista arvioitiin neuropsykologisella tutkimuksella, mielialakyselyllä ja aivojen rakenteellisella ja toiminnallisella MRI (fMRI) -tutkimuksella akuuttivaiheesta aina 6 kk:n vaiheeseen. Äänikirjoihin verrattuna musiikin kuuntelu edisti puhetoimintojen ja kielellisen muistin kuntoutumista sekä vähensi negatiivista mielialaa. Laulumusiikilla oli voimakkain vaikutus sekä puheen että muistin kuntoutumiseen etenkin afasiapotilailla. VBM- ja fMRI-tulokset osoittivat, että laulumusiikin kuuntelu lisäsi harmaan aineen tilavuutta vasemmalla ohimolohkolla ja toiminnallista konnektiivisuutta oletustilaverkostossa 6 kk aikana.

Tämän väitöskirjan tulokset tuovat lisää näyttöä päivittäisen musiikin kuuntelun positiivisesta vaikutuksesta ja tukevat sen käyttöä toimivana, helppona ja edullisena työkaluna, joka edistää AVH:n jälkeistä kognitiivista ja emotionaalista toipumista. Tämä juontuu rakenteellisista ja toiminnallisista muutoksista ohimo- ja päälakilohkoalueiden verkostoissa, mitkä ovat ratkaisevia tunteiden käsittelyn, kielen sekä muistin kannalta. Tutkimus tuo uutta tietoa etenkin laulumusiikin kuuntelun vaikutuksesta kuntoutumiseen sekä laulujen käytöstä oppimisen ja muistin tukena, erityisesti afasiasta toipumisessa.

Acknowledgements

This thesis is a fruitful outcome of collaboration between the Cognitive Brain Research Unit (CBRU) at the Department of Psychology and Logopedics, Faculty of Medicine, University of Helsinki; the Division of Clinical Neurosciences, Turku University Hospital (TYKS), and the Medical Imaging Centre of Southwest Finland. This study was financially supported by Finnish Graduate School of Psychology, Doctoral school of Psychology, Learning and Communication, Miina Sillanpää Foundation, Finnish Cultural Foundation, and Academy of Finland (project funding and funding for the Finnish Centre of Excellence in Interdisciplinary Music Research led by Professor Petri Toiviainen 2008-2013).

Most of all, I am extremely grateful to my supervisors: Associate Professor Teppo Särkämö, Research director Mari Tervaniemi, and Professor Seppo Soinila for all the guidance, support and advice for my work, and also for all the interesting and fruitful discussions about music, neuroscience, neurology and life itself we have had over the years. I consider myself highly privileged for having the possibility to learn from you. I thank you Teppo for your patience, encouragement, and for being the mainstay of this study. I admire your determination, clear vision and the courage. I thank you Mari for the warm support and guidance, and being the safe haven of this study during these years. I thank you Seppo for the neurological expertise, practical help in recruiting the patients and organizing the practical issues in TYKS and for shared enthusiasm for music.

This study would not have been possible without expertise and enthusiasm I have experienced when planning and implementing this study. I want to express my deepest gratitude to our wonderful radiographers Ulla Anttalainen (†), Tuija Vahtera, and Riku-Pekka Luoto, and our fantastic music therapists Terhi Lehtovaara, Pekka Rajanaro, and Aki Ylönen. I also want to express my humble gratitude to the personnel in the Acute Stroke Unit and Department of Neurology in TYKS, and to all of the participants of the study and their families. In addition, I want to thank Professors Matti Laine, Riitta Parkkola, Minna Huotilainen, Petri Toiviainen, and Antoni Rodríguez Fornells as well as Assistant Professor Pablo Ripollés, Dr. Jani Saunavaara, Dr. Jyrki Tuomainen, and music therapist Sari Laitinen for their assistance and expertise.

I was very fortunate to receive reviews of my thesis from two esteemed experts, Professor Sonja Kotz and Professor Heikki Hämäläinen. I am thankful for their valuable comments and encouraging and positive evaluations.

I want to express my special gratitude to my research colleague Dr. Aleksi Sihvonen for support and encouragement and for sharing this journey with me, with all the data collection, which included tens of canned saliva samples, reeled miles of the fMRI cable, and hours of discussions, but most of all for friendship. My co-writer Dr. Tanja Linnavalli deserves special thanks for friendship and help, but also for lending her voice to the sung-spoken story recall task and being a key person in creating it. Laboratory engineer Tommi Makkonen deserves special thanks for his priceless help and always friendly support with technical matters. Furthermore, I want to thank the brilliant people at CBRU: Dr. Paula Virtala, Dr. Ritva Torppa, Anni Pitkäniemi, Emmi Pentikäinen, Sini-Tuuli Siponkoski, Nella Moisseinen, Dr. Eino Partanen, Katri Saarikivi, Miika Leminen and all the other talented people in the unit.

I have been fortunate having the possibility to work as a clinician in the Masku Neurological Rehabilitation Centre. I thank the personnel of rehabilitation center for supporting atmosphere and letting me to bring my ideas to the clinical work. I thank especially Docent Päivi Hämäläinen, neuropsychologist Emilia Häyry, and speech therapist Piia Siivonen for warm support and friendship.

I express my special gratitude to my dearest and nearest; you are my essential roots. Thank you, my dearest friends Kati Urho, Janne Merisaari, Johanna Litzen and Nicola Guerra, and Anna Pelli for all the discussions, laugh, and company over the years. Many relatives and friends have played an important role in my life and should be mentioned here, I hope you recognize yourself. I am incredibly grateful to my parents Pirjo (1949-2020) and Jukka for guiding me with love throughout my life and for teaching me that only sky is the limit if I am ready to work enough for my goals. My mother's calming motto "everything is going to work out, in the way or another, you'll see", has comforted me uncountable times. I thank my sister Nora and niece Naomi for love and encouragement. I am grateful to my in-laws Erja and Ilpo Murtojärvi for warmly taking me as part of your family, and also to my sisters- and brothers-in-law, Saana, Aleksi and Marjukka, Julius and Sarita for encouragement, friendship and shared enthusiasm for food. Finally, I owe my deepest gratitude to my family: my husband Ville Murtojärvi for all his love, sense of humor, support, and understanding. You are my solid rock and real-life hero. Casper, Carla, and Linnea, dear apples of my eye, I am extremely grateful to have you. I thank you for all the flexibility this study has demanded from you. You give me constant joy and remind me of the priorities of life and of where I belong.

Lastly, thank you for the music! It connects us and makes life's ups more enjoyable and downs more bearable.

Turku, August 2020

Vera Leo

Contents

List of original publications

I Leo, V., Sihvonen, A. J., Linnavalli, T., Tervaniemi, M., Laine, M., Soinila, S., & Särkämö, T. (2018). Sung melody enhances verbal learning and recall after stroke. *Annals of the New York Academy of Sciences,* 1423, 1, 296-307.

II Leo, V., Sihvonen, A. J., Linnavalli, T., Tervaniemi, M., Laine, M., Soinila, S., & Särkämö, T. (2019). Cognitive and neural mechanisms underlying the mnemonic effect of songs after stroke. *NeuroImage: Clinical,* 24, 101948.

III Sihvonen, A.J.*, Leo, V.*, Ripollés, P., Lehtovaara, T., Ylönen, A., Rajanaro, P., Laitinen, S., Forsblom, A., Saunavaara J., Parkkola, R., Autti, T., Silvennoinen, H. M., Laine, M., Rodríquez-Fornells, A., Tervaniemi, M., Soinila, S., & Särkämö, T. (*in revision*). Vocal music enhances language and memory recovery after stroke: Pooled results from two RCTs.

*Aleksi J. Sihvonen and Vera Leo contributed equally to the work (shared 1st authorship).

Abbreviations

ABG	audiobook group
AF	arcuate fasciculus
ANOVA	analysis of variance
CG	central gyrus
FDR	false discovery rate
fMRI	functional magnetic resonance imaging
FWE	family-wise error
GMV	grey matter volume
IFOF	inferior fronto-occipital fasciculus
ILF	inferior longitudinal fasciculus
IMG	instrumental music group
IPL	inferior parietal lobule
ITG	inferior temporal gyrus
MBEA	Montreal Battery of Evaluation of Amusia
MG	music group
MNI	Montreal Neurological Institute
MOG	middle occipital gyrus
MRI	magnetic resonance imaging
MTG	middle temporal gyrus
NIHSS	National Institute of Health Stroke Scale
PHG	parahippocampal gyrus
PE	primacy effect
POMS	Profile of Mood States
RBMT	Rivermead Behavioural Memory Test
RE	recency effect
RT	reaction time
SD	standard deviation
SEM	standard error of the mean
SFG	superior frontal gyrus
SPE	serial position effect
SSSRT	Sung-Spoken Story Recall Task
STG	superior temporal gyrus
TR	repetition time
UF	uncinate fasciculus
VMG	vocal music group
WMV	white matter volume

1 INTRODUCTION

Cerebrovascular disease, such as stroke, affecting the brain is worldwide the second leading cause of death and the third leading cause of disability (Global Health Estimates. Geneva: World Health Organization, 2012). Approximately 80% of strokes are ischemic strokes, caused by a blockage in one or more arteries leading to the brain (Langhorne, Bernhardt & Kwakkel, 2011). About 15 % of strokes are hemorrhagic strokes, caused by a ruptured blood vessel that bleeds into brain tissue. Five percent of strokes are unspecified (Langhorne et al., 2011). The effects of a stroke depend on which part of the brain is injured, how extensive is the lesioned area, and how severe are the symptoms. The prevalence of stroke has been increasing (Feigin et al., 2014) due to the growing number of the aging population. The increased prevalence of stroke entails an enormous economic and societal burden (DiLuca & Olesen, 2014) and this has raised the need for motivating and effective rehabilitation tools.

During the last decades, the scientific and clinical interest towards the potential of music as a rehabilitative tool to facilitate recovery of neurological illnesses has been steadily growing (Magee, Clark, Tamplin & Bradt, 2017; Sihvonen et al., 2017a). Music, beside spoken language, has been a key part of every known human culture throughout history. Music is an important source of enjoyment and well-being across life and furthermore a versatile and powerful stimulus for the brain. The development of modern brain imaging methods enables us to discover and understand better what kind of effects music generates in humans and how it can be used effectively as a neurological rehabilitation tool to facilitate recovery and rehabilitation, and promote well-being (Särkämö, Tervaniemi & Huotilainen, 2013; Sihvonen et al., 2017a).

1.1 Stroke in Finland

In Finland, the population is aging rapidly, and cerebrovascular diseases affecting the brain are becoming more and more common among the elderly population, meaning altogether 25 000 new cases every year (Aivoliitto, 2013). Of all cerebrovascular diseases, stroke is the most common. About 25 % of stroke survivors recover fully and over 50 % recover to live independently (Aivoliitto, 2013). However, 50 % of stroke survivors have a permanent neurological impairment and disability, 25 % of them have a disability that remains severe, and approximately 25 % pass away within a year from stroke (Aivoliitto, 2013). Even though the development of acute medical care for stroke (e.g., thrombolysis) has improved its long-term prognosis, stroke still causes more permanent severe disability than any other neurological disorder. In addition, because they are so common and often require long periods of institutional care, cerebrovascular diseases are the third most expensive disease in Finland after mental disorders and dementia. In its entirety, cerebrovascular diseases incur 7% of the total costs of Finnish healthcare (Aivoliitto, 2013).

1.1.1 Cognitive implications of stroke

The most common impairments caused by stroke are motor and somatosensory deficits of the upper and lower extremities, affecting around 80% of patients (Rathore, Hinn, Cooper, Tyroler & Rosamond, 2002), but also cognitive symptoms are relatively common. Among the cognitive symptoms, stroke patients most frequently experience deficits in executive functioning (32-39%), visual perception and construction (32-38%), reasoning (24-26%) as well as language and verbal memory (22-26%) (Nys, van Zandvoort, de Kort & Jansen, 2005; Nys et al., 2007). In addition, about 15-42% of stroke survivors have aphasia (Inatomi et al., 2008; Kadojic et al., 2012; Engelter et al., 2006; Flowers et al., 2013), a deficit in speech production and/or comprehension, and 35-69% have acquired amusia (Särkämö et al., 2009; Ayotte, Peretz, Rousseau, Bard & Bojanowski, 2000; Schuppert, Münte, Wieringa, & Altenmüller, 2000), a deficit in music perception or production. Aphasia and amusia are naturally associated with the impaired processing of speech and melodies, yet the neural processing of vocal music seems to be relatively well preserved in these conditions after stroke (Sihvonen et al., 2017b). Studies suggest a clear co-morbidity between aphasia and amusia after stroke, with 40-50% of persons with amusic persons also having at least mild aphasia (Särkämö et al., 2009; Sihvonen et al., 2016; Stewart, von Kriegstein, Warren & Griffiths, 2006), even though the two disorders can occur also independently and separately (Peretz & Coltheart, 2003).

1.1.2 Rehabilitation after stroke

Most of the spontaneous recovery process occurs during the first six months post-stroke, when recovery-related neuroplasticity, meaning the capacity of the brain to change and adapt after injury, is at its' highest. Spared brain regions and networks in both ipsi- and contralesional hemispheres show both structural and functional neuroplasticity changes (Cramer, 2008) and contribute to the recovery of cognitive and motor functions (Tang et al., 2015; Umarova et al., 2016), and language (Saur et al., 2006; Hartwigsen & Saur, 2019). According to the 2016 Current Care Guidelines for Stroke, which is the Finnish evidence-based clinical practice guideline for stroke, at the acute stage the treatment concentrates on restraining the occurred damage, saving the endangered brain tissue, and alleviating the somatic symptoms related to the stroke. Once the physiological or neurological state of the patient is stable enough, the treatment then involves rehabilitation, the need and implementation of which are evaluated individually (Stroke: Current Care Guidelines, 2016). Early started and intensive rehabilitation is crucial in the recovery of motor, cognitive, and emotional impairment caused by stroke and for improving the quality of life and everyday functioning of the patient (Dobkin, 2004; Diserens & Rothacher, 2005; Langhorne et al., 2011).

In the current social and economic situation, the availability of rehabilitation services for stroke patients is uncertain, especially regarding services targeted for the rehabilitation of cognitive impairments. In practice, this often leads to patients receiving rehabilitation that is suboptimal in its quantity (too little), frequency (too sparsely spaced), and timing (too late), especially from the view of optimal plasticity of the brain. Alarmingly, it has been noted internationally that in rehabilitation units stroke patients spend most of their time (50-70%) unengaged in any activity, mostly lying in the bed and without social interaction (Bernhardt, Dewey, Thrift & Donnan, 2004; De Wit et al., 2005). Typically, stroke patients have guided rehabilitative activity approximately 20% per day (Mackey, Ada, Heard & Adams, 1996), which is insufficient for many.

Considering the evidence for the stimulating role of the post-stroke environment and activities on functional recovery and brain plasticity (Nithianantharajah & Hannan, 2006; Cramer, 2008), and the fact that the current rehabilitation services do not meet the true needs of the patients, it is important to develop new rehabilitation tools. The rapidly growing prevalence of stroke in the ageing population (Feigin et al., 2014) has led to increasing need especially for effective and motivating rehabilitative tools that self- or caregiver-applicable (can be used by the patients themselves or with the aid of family members) and therefore rely less on the rehabilitation resources of the public health care system. Recent studies performed at stroke units have demonstrated that implementing an environmental enrichment intervention increases physical, social, and cognitive activity levels (Janssen et al., 2014; Rosbergen et al., 2017) and also enhances mood and general cognitive status (Khan et al., 2016). In the environmental enrichment intervention, stroke patients are provided with additional social interaction and stimulating equipment and activities such as iPads, puzzles, games and music (Janssen et al., 2014; Rosbergen et al., 2017). By definition, environmental enrichment provides sensory, motor, social, and cognitive stimulation to increase overall activity in the environment (Nithianantharajah & Hannan, 2006). Environmental enrichment has been shown, in animal studies of stroke, to be a powerful driver of neuroplasticity, enhancing dendritic branching, neurotrophin production, and neuro- and angiogenesis, and also potentially improving functional recovery (Johansson, 2004; Livingston-Thomas et al., 2016), also in the domain of learning and memory (Dahlqvist, Rönnbäck, Bergström, Söderström & Olsson, 2004; Gobbo & O'Mara, 2004).

As a form of auditory environmental enrichment, music provides a multidomain stimulus that is both pleasant and rewarding and engages the brain extensively, and can thereby be used to facilitate activity-dependent neuroplasticity in the large-scale brain network that it stimulates (Särkämö and Soto, 2012).

1.2 Music and speech processing in healthy brain

1.2.1 Neural basis of music and speech perception

Functional neuroimaging studies in healthy subjects have revealed that music and speech share processing resources in many parts of the brain but also engage partly separate regions and show a degree of hemispheric lateralization (LaCroix, Diaz, & Rogalsky, 2015). At the auditory perceptual level, they both involve the encoding and analysis of the spectrotemporal acoustic structure of sounds, which takes place along the ascending auditory pathway from the ear to the auditory cortex in the superior temporal lobe (Zatorre, Belin & Penhune, 2002; Hickok & Poeppel, 2007). The auditory cortex is more specialized on the fine-grained spectral processing of sounds in the right hemisphere and on the rapid temporal processing of sounds in the left hemisphere (Zatorre et al., 2002; Tervaniemi & Hugdahl, 2003).

After this acoustic analysis, speech perception entails a sublexical (phonological) analysis, which engages specifically the left superior temporal sulcus (Turkeltaub & Coslett, 2010), followed by mapping the phonological representations onto articulatory-motor and lexical conceptual representations, which activate left frontoparietal and middle temporal regions, respectively (Hickok & Poeppel, 2007). This information is finally combined in the left inferior frontal gyrus (Broca's area), which has a critical role in syntactic and semantic processing (Patel, 2003; Rodd, Vitello, Woollams & Adank, 2015).

Music perception, in turn, features the analysis of higher-level musical features (e.g., chords, melody), which are processed by a largely bilateral (but with emphasized role of the right hemisphere) network comprising inferior and medial prefrontal areas, premotor areas, superior temporal areas, and inferior parietal areas (Janata et al., 2002; Patel, 2003; Alluri et al., 2012). The rhythm of music additionally recruits the motor and premotor cortices, basal ganglia, and cerebellum (Grahn & Rowe, 2009; Zatorre, Chen & Penhune, 2007). Also working memory and episodic memory functions in multiple prefrontal cortex, temporal, parietal, and subcortical areas are engaged when keeping music in mind and in recalling memories associated with music (Janata, 2009; Jerde, Childs, Handy, Nagode & Pardo, 2011). Finally, the emotional impact of music and the hedonic pleasure derived from it is closely linked to the mesolimbic reward network of the brain, involving especially the nucleus accumbens, amygdala, hippocampus, cingulate cortex, and orbitofrontal cortex (Koelsch, 2014; Zatorre & Salimpoor, 2013).

1.2.2 Speech and music combined: songs and singing in the brain

Songs bind lyrics and melody into a unified representation and therefore provide an interesting interface between speech and music. The neural processing of songs is a combination of activation patterns for the processing of linguistic (syntactic and semantic), musical (melodic and rhythmic), domain-general cognitive (attention and memory), vocal-motor, and emotional information (Zarate, 2013). Compared to speech or instrumental music, listening to vocal music (songs with lyrics) activates more bilateral temporal [e.g., superior temporal gyrus (STG), planum temporale] and frontal [e.g., superior frontal gyrus (SFG), premotor cortex] areas as well as subcortical / limbic areas (e.g., hippocampus, striatum, orbitofrontal cortex) more extensively than listening to speech (Callan et al., 2006; Schön et al., 2010) or instrumental music (Alluri et al., 2013; Brattico et al., 2011). This suggests that the perception of singing engenders stronger and more wide-spread auditory, linguistic, and emotional processing in the brain than speech or music alone. The perception of vocal melodies also engages the right sensorimotor cortex and its connectivity with other parts of the auditory dorsal stream, a key pathway of audio-vocal integration in speech production (Lévêque and Schön, 2015).

The production of singing, in turn, involves continuous interaction between two cortical systems, the dorsal vocal production pathway from parietal to frontal regions and the ventral auditory perception pathway from temporal to frontal regions (Zarate, 2013). Additionally, also the abovementioned wide-scale language and music processing networks are closely linked to singing (Brown, Martinez, Hodges, Fox, & Parsons, 2004; Callan et al., 2006; Özdemir, Norton, & Schlaug, 2006; Kleber, Veit, Birbaumer, Gruzelier, & Lotze, 2010; Zarate, 2013). Compared to speaking, singing activates, albeit stronger in right than left, bilateral superior temporal areas, pre- and post-central gyri, and inferior frontal areas more extensively (Brown et al., 2004; Callan et al., 2006; Özdemir et al., 2006).

1.3 Music as a verbal learning aid

1.3.1 Mnemonic effect of songs in healthy subjects and in neurological patients

The role of music as an enhancer and learning aid for verbal memory and recall, a phenomenon common in everyday life for example in children's rhymes and learning songs, such as the Alphabet song, has been under scientific interest for decades. Musical mnemonics is a term sometimes used to denote that singing, but music also in general, provides a structured temporal scaffolding framework that facilitates word learning and therefore acts as a learning and mnemonic aid (Ferreri & Verga, 2016). Experimental studies have revealed that (i) hearing the melody of a well-known song can cue the retrieval of its lyrics (Rubin, 1977) and that (ii) lyrics are effectively paired with the melody when learning unfamiliar songs (Crowder, Serafine & Repp, 1990; Samson & Zatorre, 1991). Likewise, in comparing memory performance for novel verbal material presented in sung or spoken formats, studies have shown better learning and / or delayed recall of sung material (Calvert & Tart, 1993; Ludke, Ferreira, & Overy, 2014; McElhinney & Annett, 1996; Rainey & Larsen, 2002; Tamminen, Rastle, Darby, Lucas, & Williamson, 2017; Wallace, 1994). The same phenomenon has been found also in implicit learning: in implicit learning paradigms, the pre-attentive segmentation and learning of novel trisyllabic words has been shown to be enhanced when presented in sung format in adults, or in the prosodically rich infant-directed speech in infants (Bosseler, Teinonen, Tervaniemi, & Huotilainen, 2016; Schön et al., 2008).

In most studies, the enhancement of memory for sung novel material has been reported in the learning and/or delayed recall of connected text, such as song lyrics (Wallace, 1994; McElhinney & Annett, 1996; Kilgour, Jakobson & Cuddy, 2000; Purnell-Webb & Speelman, 2008), slogans (Yalch, 1991), and sentences (Calvert & Tart, 1993), while in some studies, it has been reported in the learning or recall of unconnected text, such as word lists (Chazin & Neuschatz, 1990; Rainey & Larsen, 2002). However, in some studies no advantage has been found in the direct comparison sung and spoken presentation formats (Racette & Peretz, 2007; Thaut, Peterson, & McIntosh, 2005) or when controlling for differences in presentation rate, which is typically around 50% slower in singing than in speech (Kilgour, et al., 2000). For example, in the Racette et al. (2007) study, the sung presentation of unfamiliar lyrics showed no advantage when it was compared to a spoken presentation that had the same overall duration and was coupled with the melody sung on syllable /la/.

In summary, it seems that the learning of novel verbal material is enhanced by sung presentation, but this effect seems to be dependent on some conditions. First, the effect emerges more robust for verbal material comprising multiple verses of connected text, in which individual words are linked in a meaningful way, than for unconnected text. Second, the sung melody needs to be relatively simple and repetitive across the verses, drawing attention to the surface characteristics of the text, such as phonemic and rhythmical properties and phrasing that are accentuated by the melody, making it easier to learn (Wallace, 1994). Third, the singing must be natural in terms of tempo, meaning the speed in which words are pronounced, as speeding up the sung version to match the duration of the spoken version abolishes the effect (Kilgour et al., 2000). Overall, the melodic and rhythmic patterns of music provide a rich structure that can potentially help in linking words and phrases and identifying line lengths and stress patterns, functioning as a verbal memory aid (Wallace, 1994). Clinically, this could have implications for facilitating verbal learning especially in patients with language or cognitive (especially memory) impairments due to neurological illness.

In clinical populations, the memory advantage of sung over spoken verbal material has previously been observed in the learning and recall of word lists in patients with multiple sclerosis (Thaut, Peterson, McIntosh & Hoemberg, 2014). In addition, the advantage of the sung material has been found in the recognition of unfamiliar song lyrics in patients with Alzheimer's disease (Simmons-Stern, Budson & Ally, 2010; Moussard, Bigand, Belleville & Peretz 2014; Palisson et al., 2015) and in patients with amnesia (Haslam & Cook, 2002). Also patients with aphasia have been observed to repeat the words of familiar songs (Straube, Schulz, Geipel, Mentzel, & Miltner, 2008) as well as complete the ends of phrases of familiar songs (Kasdan & Kiran, 2018) better when they are presented in sung than spoken format. However, none of the previous aphasia studies have reported an advantage of sung over spoken presentation in the learning of unfamiliar lyrics (Hébert, Racette, Gagnon, & Peretz, 2003; Racette, Bard, & Peretz, 2006; Straube et al., 2008), and thus the mnemonic impact of songs in aphasia is still unknown.

1.3.2 Memory mechanisms underlying the mnemonic effect of songs

The memory mechanisms underlying the advantage of sung over spoken presentation are still largely unexplored. In addition, the extent to which the processing of lyrics and melody in songs is dissociated or integrated in memory is still under debate. Older lesion studies suggest a dissociation between memory for lyrics and melodies (Samson & Zatorre, 1991; Hérbert & Perez, 2001). In contrast, recent behavioral and fMRI studies have reported more integrated processing of lyrics and tunes across different frontotemporal regions, especially in the left hemisphere (Fedorenko, Patel, Casasanto, Winawer & Gibson, 2009; Sammler et al., 2010; Schön et al. 2010). If the two are integrated in memory, then the melody can simply provide an associative memory cue for the verbal content, thereby increasing the likelihood of better recall.

Another potential memory mechanism underlying the mnemonic effect of songs is the *serial position effect* (SPE). The SPE is a classic learning-related phenomenon whereby information presented early and late in a sequence is remembered better than information presented in the middle of the sequence (Crowder & Greene 2000; Deese & Kaufman 1957). Commonly, in a learning task where the number of serially presented items overruns working memory capacity, the SPE is reflected by a Ushaped curve with an increased likelihood to recall the first items (*primacy effect*, PE) and the last items (*recency effect,* RE) compared to the ones in the middle. The classic dual-storage memory model (Atkinson & Shiffrin, 1968) qualifies the RE to immediate access from a short-term storage buffer and the PE to retrieval from a long-term memory storage through working memory. In neuroimaging studies, there is some evidence that the PE is associated with dorsolateral prefrontal areas and the RE with inferior parietal, temporal, and hippocampal areas (Buchsbaum, Ye, & D'Esposito, 2011; Düzel, Hufnagel, Helmstaedter, & Elger, 1996; Innocenti et al., 2013; Spalletta et al., 2016; Staffaroni et al., 2017). It is possible that the learning and recall of sung verbal material could show a different pattern of SPEs compared to spoken verbal material, but experimental evidence supporting this is limited and concerns only healthy persons. Silverman (2007) found a significant interaction between serial position and condition, with a less bowed U curve indicating more stable recall performance across the list, especially for the middle items in the sung condition, when comparing the immediate recall of digit sequences presented in spoken and sung format in healthy subjects. In Maylor's (2002) study, it was found that the degree of familiarity of song lyrics was linked to a better recall across all verses of the song, albeit both PE and RE were also observed.

Chunking, a process of grouping individual items into larger units, is another well-known memory mechanism that can facilitate learning and optimize performance by decreasing memory load (Bor, Duncan, Wiseman, & Owen, 2003; Gobet et al., 2001). Chunking-based encoding of verbal material has been linked to a wide-scale bilateral network comprising of dorsolateral prefrontal, inferior parietal and posterior temporal regions (Bor, Cumming, Scott, & Owen, 2004). Chunking training in working memory has been observed to enhance language processing in aphasia (Eom & Sung, 2016). Wallace (1994) has suggested that the repetitive melodic and rhythmic structure of songs can serve as an encoding and retrieval cue by chunking consecutive items into melodic phrases and assisting in positioning and sequencing textual units, thus decreasing the likelihood that units will be misplaced and disrupt memory for succeeding units in a sung verbal learning task. Yet there is currently very little experimental evidence on chunking in the context of music or singing, with only two studies in healthy subjects reporting greater chunking of recalled material for words presented with background music vs. silence (Ferreri, Bigand, Bard, & Bugaiska, 2015) and via singing vs. speaking (McElhinney & Annett, 1996).

1.4 Music in neurological rehabilitation

Converging evidence from neuroimaging studies showing that music processing is underpinned by wide-scale bilateral brain networks (see section 1.2.1), which show both structural and functional neuroplasticity changes after regular musical training or activity (Herholz & Zatorre, 2012), has motivated the development a number of therapist-led music interventions utilizing rhythmic entrainment to music, playing musical instruments, and singing. Overall, these interventions have shown to be effective in improving the recovery of gait (Schauer & Mauritz, 2003; Thaut et al., 2007), upper-extremity motor functions (Schneider, Münte, Rodriguez-Fornells, Sailer & Altenmüller, 2010; van Vugt et al., 2016), and speech production (van der Meulen, van de Sandt-Koenderman, Heijenbrok-Kal, Visch-Brink & Ribbers, 2014; Raglio et al., 2016). Importantly, active music interventions have also been demonstrated to induce recovery-related neuroplasticity changes in auditory-motor (Altenmüller, Marco-Pallares, Münte, & Schneider, 2009; Grau-Sánchez et al., 2013; Ripollés et al., 2016) and speech networks (Schlaug, Marchina & Norton, 2008; Jungblut, Huber, Mais & Schnitker, 2014; Wan, Zheng, Marchina, Norton & Schlaug, 2014). The following two sections will focus on music-based interventions for stroke utilizing singing and music listening.

1.4.1 Singing-based stroke rehabilitation

The majority of research on singing-based rehabilitation in aphasic stroke patients has utilized a method called *melodic intonation therapy*. Melodic intonation therapy is a classic rehabilitation method of non-fluent aphasia that was developed based on the observation that persons with even severe aphasia can often produce linguistically accurate and well-articulated words while singing, but not when speaking (e.g. Albert, Sparks & Helm 1973; Schlaug, Marchina & Norton, 2009). Melodic intonation therapy is a hierarchically structured treatment method that uses sung patterns to magnify the normal melodic content of speech by translating spoken phrases into melodically intoned patterns using two or more pitches, which create usually the interval of a minor third (three semitones). It provides a good approximation of the prosody of speech that still falls into the category of singing (Norton, Zipse, Marchina & Schlaug, 2009; Schlaug et al., 2008). Melodic intonation therapy is designed to lead patients with non-fluent aphasia from singning simple 2-3 syllable phrases to speaking phrases of five or more syllables across three levels of treatment. The core elements of melodic intonation therapy are the inherent continuous melodic intonation and rhythmic tapping of each syllable while phrases are intoned and repeated. Melodic intonation therapy is an intensive therapy, for it typically involves training 1.5 h/day, 5 days/week until the patient has mastered all three levels (Helm-Estabrooks & Albert, 2004; Norton et al., 2009; Albert et al., 1973; Schlaug et al. 2009).

Previous behavioral studies suggests that melodic intonation therapy has a positive impact to speech recovery, such as in spontaneous speech output, articulation, and naming ability, in patients with non-fluent aphasia (e.g. Belin et al., 1996; Wilson, Parsons & Reutens, 2006; Schlaug et al., 2008, 2009; Bonakdarpour, Eftekharzadeh & Ashayeri, 2003; van der Meulen et al., 2014) and particularly for patients with large left-hemispheric lesions (Schlaug et al., 2008). However, most of the studies concentrate on the chronic phase of post-stroke or are single case studies (van der Meulen et al., 2014). There are also studies that have failed to demonstrate the superiority of the singing over speech in non-fluent aphasia (Zumbansen, Peretz & Hébert, 2014a; Hébert et al., 2003; Peretz, Gagnon, Hébert & Macoir, 2004).

A number of neural, cognitive, and emotional mechanisms underlying the efficacy of melodic intonation therapy have been proposed, although experimental research testing them is still scarce. Some studies report that the key mechanism of melodic intonation therapy is the use of the melodic element that engages the right hemisphere vocal-motor and auditory regions coupled with simultaneous rhythmic tapping of the left hand. That in turn engages potentially the right-hemispheric sensorimotor network that coordinates hand movements, but also orofacial and articulatory movements (Schlaug et al., 2008, 2009; van der Meulen et al., 2104). In addition, using melody and accentuated prosodic features leads to general reduction in the vocalization rate as syllables are lengthened and chunked into larger structures. Similarly, once the right temporal lobe is engaged by the melodic intonation and contour, the role of the left hand tapping is likely to be the activation and priming of a right-hemispheric sensorimotor network for articulation (Schlaug et al., 2009). Some studies, in contrast, have suggested that the efficacy of melodic intonation therapy lies on reactivating essential motor language zones, like the Broca's area, in the left hemisphere, while reducing abnormal activations in the right hemisphere (Belin et al., 1996; Laine, Tuomainen & Ahonen, 1993a; Breier, Randle, Maher & Papanicolaou, 2010). In addition to the neuroplastic reorganization of the speech network, other proposed mechanisms include activation of the mirror neuron system and multimodal integration, utilization of shared or specific features of music and language, and motivation and mood (Merrett, Peretz & Wilson, 2014).

1.4.2 Music listening as a rehabilitation tool in stroke

Listening to music is seen as a safe and easily accessible intervention to facilitate stroke recovery, already from the early (acute) post-stroke stage when active rehabilitation methods are often not feasible. Särkämö and colleagues (Särkämö et al., 2008, 2010, 2014; Forsblom, Laitinen, Särkämö & Tervaniemi, 2009) were first to explore the long-term effects of daily music listening on stroke recovery. In a three-arm randomized controlled trial, stroke patients (N = 55) either listened daily (min 1 hour per day) to self-selected music or audiobooks using a portable player or received standard care only during 2 months after an acute middle cerebral artery stroke. The recovery was assessed at acute, 3-month, and 6-month stages. Music listening was found to improve the recovery of verbal memory and focused attention more than audiobook listening or standard care at 3-month and 6-month stages, and also to reduce self-reported depression and confusion more than standard care at the 3-month stage (Särkämö et al., 2008). Using voxel-based morphometry, listening to music was also found to increase grey matter volume in left and right SFG, left anterior cingulate, and right ventral striatum compared to audiobooks and standard care in left hemisphere-lesioned patients (Särkämö et al., 2014). Subjectively, the patients reported that music helped them to relax and increased positive mood and motor activity more than audiobooks (Forsblom et al., 2009).

The approach in the study of Särkämö et al. was to make the listening intervention as naturalistic as possible and therefore the patients were able to listen to any genre or type of music they wished. The music selections of the patients were diverse, both across and within subjects, ranging from popular music to jazz, folk, and classical music, with around 60% of the selected musical pieces containing lyrics (Särkämö et al., 2008), as musical preferences are typically very individual (Rentfrow, Goldberg & Levitin, 2011). Also, in a more recent randomized controlled trial study of Baylan and colleagues (2018, 2019), comparing music listening with or without additional mindfulness training to audiobook listening, music listening was reported to improve verbal memory more than audiobooks over 6 months (Baylan et al., 2019). Qualitatively, music listening was most strongly associated with increased activity, memory reminiscence, and improved mood (Baylan et al., 2018). Together, these results indicate that during the first months after an acute stroke, daily music listening can potentially have long-term positive effects on cognitive, emotional, and neural recovery.

One essential question, which was not addressed by the original 2008 study by Särkämö and colleagues, is which attributes of the music stimuli are specifically related to the rehabilitative effects of music listening. Listening to familiar songs at the acute stage after stroke seems to activate bilateral temporal, insular, and motor cortical areas, extending also to inferior frontal, medial parietal, and subcortical areas, more extensively than listening to instrumental versions of same songs (Sihvonen et al., 2017b). Because of the large-scale bilateral activation associated with the processing of vocal music, listening to vocal music could be more effective in enhancing the recovery after stroke than listening to speech that engages more the left hemisphere or to instrumental music that engages more the right hemisphere (Zatorre et al., 2002; Tervaniemi & Hugdahl, 2003; Rosenthal, 2016). However, knowledge of the impact of music listening during the stroke recovery process is still narrow, and more studies are needed, especially about the idiosyncrasies of music and how they mediate its rehabilitative effect.

2 AIMS OF THE STUDY

The present thesis explored the effects of vocal music after stroke, both as a potential tool to facilitate verbal learning and as a daily activity to enhance stroke recovery. Specifically, the three main aims were to

I. Explore if novel narrative verbal material (stories) is learned and recalled more effectively when presented in sung than spoken format after stroke (Study I).

II. Uncover the cognitive and neural mechanisms underlying the mnemonic effect of songs after stroke (Study II).

III. Determine the contribution of sung lyrics on the verbal, cognitive, emotional, and neural efficacy of music by comparing the effects of daily listening to vocal music, instrumental music, and audiobooks after stroke using a randomized controlled trial design (Study III).

3 METHODS

3.1 Subjects and study design

3.1.1 Studies I and II

Subjects (N = 50) were stroke patients recruited and examined between March 2013 and April 2016 at the Department of Clinical Neurosciences of the Turku University Hospital. All patients had a magnetic resonance imaging (MRI) verified acute ischemic stroke or intracerebral hemorrhage in the left or right hemisphere, primarily in middle cerebral artery territory, and at least minor cognitive impairment caused by stroke, which was verified from the medical records and by interviewing the patient and caregiving personnel. The following addition criteria were used: 1) no prior neurological or psychiatric disease, 2) no drug or alcohol abuse, 3) no prominent hearing impairment, which was verified from the medical records and by interviewing the patient, 4) no contraindications for MRI imaging, 5) age between 18-80, 6) Finnish speaking or bilingual (able to perform the neuropsychological testing battery in Finnish), and 7) able to cooperate meaning the ability to carry out the neuropsychological assessment. The recruitment of the patients was carried out as soon as possible after the hospitalization, but within three weeks post-stroke. All patients gave written informed consent. The patients underwent neuropsychological assessments and structural and functional MRI within three weeks, three months, and six months post-stroke. The baseline study was carried out approximately within 8 days from hospitalization (M= 7.83, SD= 4.71).

Of the 50 patients recruited originally for the study, two dropped out at the acute stage during baseline measurements, three dropped out before the 3-month follow-up, and one before the 6-month follow-up (due to claustrophobia, severe illness, or refusal). In addition, part of the patients were unable to perform the neuropsychological battery in its entirety, including the Sung-Spoken Story Recall Task (see below) used in Studies I and II, at the acute stage due to strong fatigue. Consequently, data from 31 patients were used in the statistical analyses of Studies I and II.

3.1.2 Study III

Subjects were stroke patients (N = 90) comprising of the Turku participants of Studies I and II (N = 50, see above) as well as part of the participants from the previous randomized controlled trial study performed in Helsinki (N = 40; Särkämö et al., 2008). The pooling of data across the two studies (referred to hereafter as Turku and Helsinki studies) was done in order to increase the sample size and statistical power in Study III. This was feasible because both the Turku and Helsinki studies had common inclusion criteria (MRI-verified acute ischemic stroke or intracerebral hemorrhage; <80 years old; right handed; able to perform the assessment in Finnish; able to co-operate; no hearing loss; no prior neurological or psychiatric disease; no substance abuse), time points of assessment (≤ 3 weeks, 3-month, and 6-month post-stroke), and outcome measures. Also, the timing, frequency, and delivery of the interventions were comparable in both studies. The studies were approved by the Ethics committees of the Hospital Districts of Southwest Finland (Turku) and Helsinki and Uusimaa (Helsinki), and performed in conformance with the Declaration of Helsinki. All patients signed an informed consent and received standard medical treatment and rehabilitation for stroke following the Current Care Guidelines for Stroke (Stroke: Current Care Guidelines, 2016). In both studies, randomization was stratified for lesion laterality (left/right) and performed as block randomization (10 blocks of three consecutive patients for left and right lesions), the order within the blocks drawn using a random number generator. A laboratory engineer not involved in the data collection generated the randomization list and the persons who performed the patient recruitment had no access to it (allocation concealment). The adherence was very good and similar between the trial sites, with 83/90 (92%) patients (Turku: 90%, Helsinki: 95%) completing the study up to the 3-month stage and 81/90 (90%) patients (Turku: 88%, Helsinki: 93%) up to the 6-month stage. The study design and participant flow of Study III is shown in Figure 1.

In the Turku study, the 50 stroke patients were randomized to three groups: vocal music group (VMG, N = 17), instrumental music group (IMG, N = 17), and audiobook group (ABG, N = 16). After the dropouts at different stages (see above), 45 patients completed the trial up to the 3-month stage and 44 up to the 6-month stage. In the Helsinki study, 60 stroke patients were originally recruited during 2004-2006 from the Department of Neurology of the Helsinki University Central Hospital. The patients were randomized to three groups (N = 20 in each): music group (MG), ABG, and control (standard care only) group. Fifty-five patients completed the trial up to the 3-month stage and 54 up to the 6-month stage. Those MG patients (N = 19) and ABG patients (N = 19) who had 3-month follow-up data were included in Study III. Based on information obtained from the listening diaries of the patients and the notes of the music therapists, the MG patients were reclassified to VMG (N=13) and IMG (N=6) depending on whether they had listened primarily (more than two-thirds of the material) to vocal or instrumental music during the intervention.

The final pooled data comprised 83 patients from baseline to 3-month stage (VMG: N=27, IMG: N=23, ABG: N=33) and 81 patients from baseline to 6-month stage (VMG: N=26, IMG: N=23, ABG: N=32). In addition the same pooled sample was used to compare the general effects of music and audiobook listening from acute stage to

3-month stage [MG (VMG and IMG combined): N= 50, ABG: N= 33] and from acute stage to 6-month stage [MG (VMG and IMG combined): N= 49, ABG: N=32].

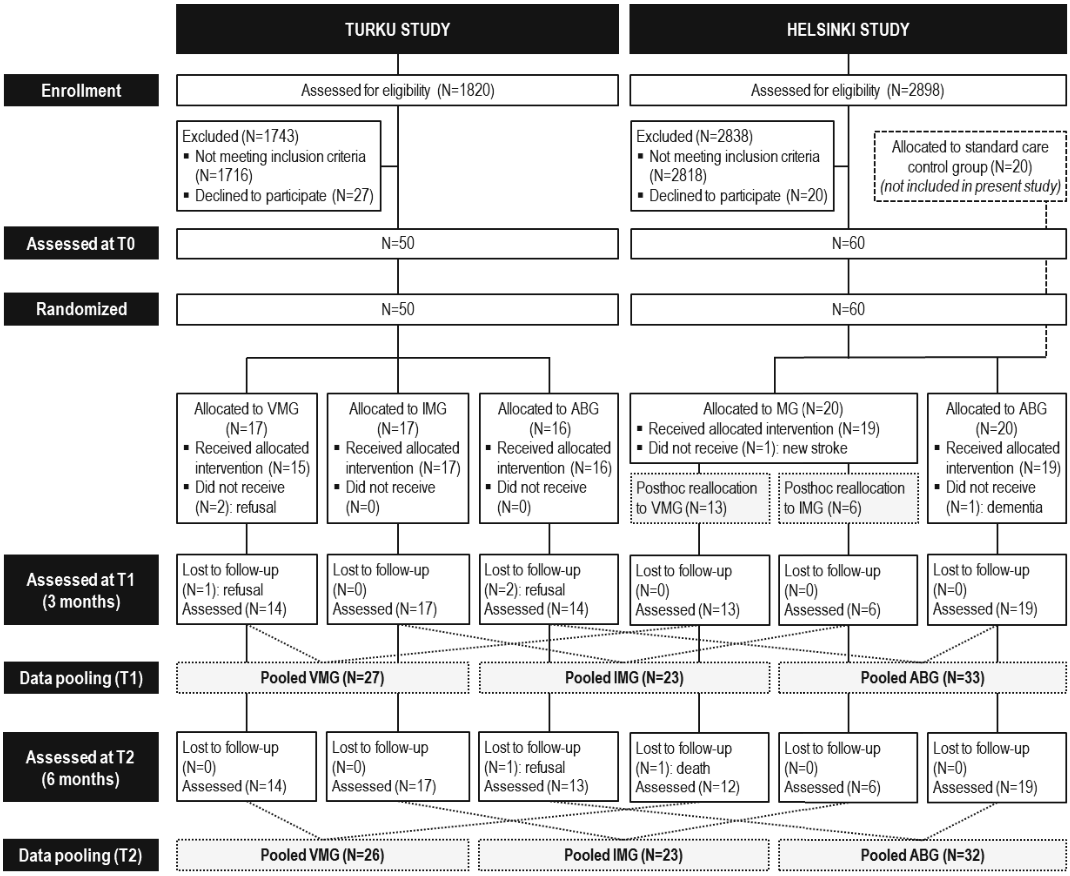

Figure 1 Flow chart outlining the design and progress of the study. Assessment for eligibility includes all neurological patients admitted to the Department of Clinical Neurosciences of Turku University Hospital (Turku study) and Department of Neurology of Helsinki University Central Hospital (Helsinki study). ABG = Audio book group, IMG = Instrumental music group, T0 = baseline (acute), T1 = 3-month stage, T2 = 6-month stage, VMG = Vocal music group. (Sihvonen & Leo et al., *in revision*)

3.2 Intervention (Study III)

After agreeing to participate in the study and performing the baseline outcome measures, a music therapist contacted the patients individually and interviewed them about their pre-stroke hobbies and leisure activities, such as music listening and reading. The music therapists informed the patients about the group allocation and provided them with a portable player and over-ear headphones and a collection of listening material. In the VMG, the material was vocal music with sung lyrics in a language that the patients understood well, mostly in Finnish or English. In the IMG, the listening material was instrumental music with no sung lyrics. In the ABG, the listening material was narrated audiobooks or radio plays with no music. Within these constraints, the listening material was selected individually with an aim to match the material to both the intervention group and the music or literature preferences of the patient as closely as possible to make the listening enjoyable. The patients were guided in using the players and they were instructed to listen to the material by themselves daily minimum one hour per day for the following two months while still in the hospital or at home. In addition, the patients were asked to keep a listening diary. During the 2-month intervention period, the music therapists kept regular contact with the patients to encourage listening, to provide more material, and to give practical aid in using the equipment if needed.

3.3 Neuropsychological assessment

Neuropsychological assessments were performed at the Department of Clinical Neurosciences in Turku University Hospital (Turku, Studies I-III) and Department of Neurology in Helsinki University Central Hospital (Helsinki, Study III) using an extensive neuropsychological testing battery, which included tests of language skills, verbal memory, focused attention, and music perception, as well as questionnaires on mood and reward value of music in life pre-stroke. The tests and questionnaires used in Studies I and II and in Study III are summarized in Table 1. In Study III, summary scores of the tests measuring the language (verbal fluency task, naming task, short Token Test), verbal memory (RBMT Story Recall, AVLT), and focused attention (CogniSpeed Stroop & Mental subtraction) domain were used in the statistical analyses in order to reduce the number of variables (Särkämö et al., 2008).

A psychologist who was blinded to the group allocation of the patients (Study III) performed the neuropsychological assessments. The assessments lasted 2-3 hours and were usually carried out in one session in a quiet room. If needed, the baseline assessment was carried out in two testing sessions to avoid interference due to fatigue. To minimize practice effects, parallel test version of the memory tests were used in different testing occasions. Reaction time (RT) tests were always performed using the better, non-paretic hand. In addition to the standardized tests and questionnaires, basic demographical and clinical information was collected and the overall severity of stroke symptoms was assessed with the National Institute of Health Stroke Scale (NIHSS) based on a neurological status examination.

Table 1. Behavioral tests and questionnaires used in Studies I-III

Domain	Measure	Task description	Studies I-II	Study III	Reference
Language skills	Verbal fluency task	List words beginning with a specific letter (60 sec)	X	X	Lezak et al. (2004)
	Naming test (CERAD / BNT)	Name items from line drawings	X	X	Morris et al. (1989) Laine et al. (1993a)
	short Token Test	Follow verbal instructions	X	X	De Renzi & Faglioni (1978)
Verbal memory	RBMT: Story recall	Remember a short story (immediate & delayed recall)	X	X	Wilson et al. (1985)
	AVLT	Remember a list of 10 words (three learning trials & delayed recall)	X	X	Lezak et al. (2004)
	SSSRT	Remember sung and spoken stories (three learning trials & delayed recall)	X		(custom task)
Focused attention	CS: Stroop	Name colors of words on screen (congruent/incongruent)		X	Revonsuo & Portin (1995)
	CS: Mental subtraction	Subtract number on screen from number 9		X	Revonsuo & Portin (1995)
Music perception	MBEA: Scale & Rhythm	Compare two consecutive melodies (same/different)	X	X	Särkämö et al. (2009) Peretz et al. (2003)
Mood	POMS	38 mood items (8 scales: Tension, Depression, Irritability, Vigor, Fatigue, Inertia, Confusion, Forgetfulness)		X	McNair et al. (1981) Hänninen (1989)
Music reward	BMRQ	20 questions about role of music in life before stroke	X	X	Mas-Herrero et al. (2013)

Measure abbreviations: AVLT = Auditory-Verbal Learning Task, BMRQ = Barcelona Music Reward Questionnaire, BNT = Boston Naming Test, CERAD = Consortium to Establish a Registry for Alzheimer's Disease, CS = CogniSpeed© test battery, MBEA = Montreal Battery of Evaluation of Amusia, POMS = Profile of Mood States, RBMT = Rivermead Behavioural Memory Test, SSSRT = Sung-Spoken Story Recall Task

3.3.1 Assessment of aphasia, amusia, and memory impairment (Studies I-III)

Aphasia was assessed using the Aphasia severity rating scale from the Boston Diagnostic Aphasia Examination (Goodglass & Kaplan, 1983). The Aphasia severity rating scale scoring was clinically done and mostly based on free conversational speech, but it also drew information from performance on standard tests of verbal comprehension, naming, and verbal fluency (see Table 1). Patients with an Aphasia severity rating scale score ≤4 were classified as aphasic (Goodglass & Kaplan, 1983).

Music perception was evaluated with a shortened version (Särkämö et al., 2009) of the Montreal Battery of Evaluation of Amusia (MBEA; Peretz, Champod, & Hyde, 2003) comprising of the Scale and Rhythm subtests (discriminating piano melodies based on melodic pitch and rhythm changes). Patients with MBEA total score under 75% correct were classified as amusic. In Study I, the patients were additionally classified as memory-impaired or not memory-impaired based on a median split of the verbal memory summary score (see above).

3.3.2 Sung-Spoken Story Recall Task (Studies I and II)

The Sung-Spoken Story Recall Task (SSSRT) is a novel, customized task, which was developed at the Cognitive Brain Research Unit to compare the learning and recall of verbal material (stories) presented in sung and spoken formats. For this purpose, sung and spoken versions of four short narrative stories (S1, S2, S3, S4) were created, which all had a common theme of an unexpected or ironic event in everyday life. The stories were on average 57 words long and all arranged in five verses. The durations of the spoken and sung versions were on average 36 and 53 seconds, respectively, the sung version being around 50% longer than the spoken version, which is a usual ratio for sung versus spoken material (Kilgour et al., 2000). A slowed-down spoken or speeded-up sung version was not used as a control in order to keep the stimuli natural sounding and the task ecologically valid and not unbearably long for the patients. All the stories were recorded by the same female voice both in spoken format, which was spoken with natural prosody, and in sung format. The sung versions of S1-2 and S3-4 had the same melodies (see Figure 2), which were composed to be simple, containing 6-7 different tones in major (S1-2) / minor (S3-4) key in 4 bars, with 4/4 meter and a tempo of 180 beats per minute (bpm). The same melody repeated in all the five verses of the song.

The spoken and sung versions of the SSSRT were presented to the patients by counter-balancing the verbal content of the stories at different stages. Thus, at the acute stage, half of the patients heard S1 spoken and S2 sung and half of them heard S1 sung and S2 spoken, and at the 6-month stage, half of the patients heard S3 spoken and S4 sung and half of them heard S3 sung and S4 spoken. At both stages, the spoken task was always presented first, with three consecutive learning trials and a delayed recall trial 25 minutes later. After a 15-minute interval, the sung task was presented following the same protocol. The fixed presentation order was chosen instead of a counter-balanced order to avoid the possibility that when performing the sung version first, the patients could then covertly use the melody for example by imagining or humming the melody in their mind while performing the spoken version. This would have been possible since the story pairs, S1-S2 and S3-S4, were designed to have a matched linguistic structure in terms of line length and phrasing so that both would work with the same melody. The stimuli were presented on a laptop computer with headphones adjusting the volume to a comfortable and clearly audible level.

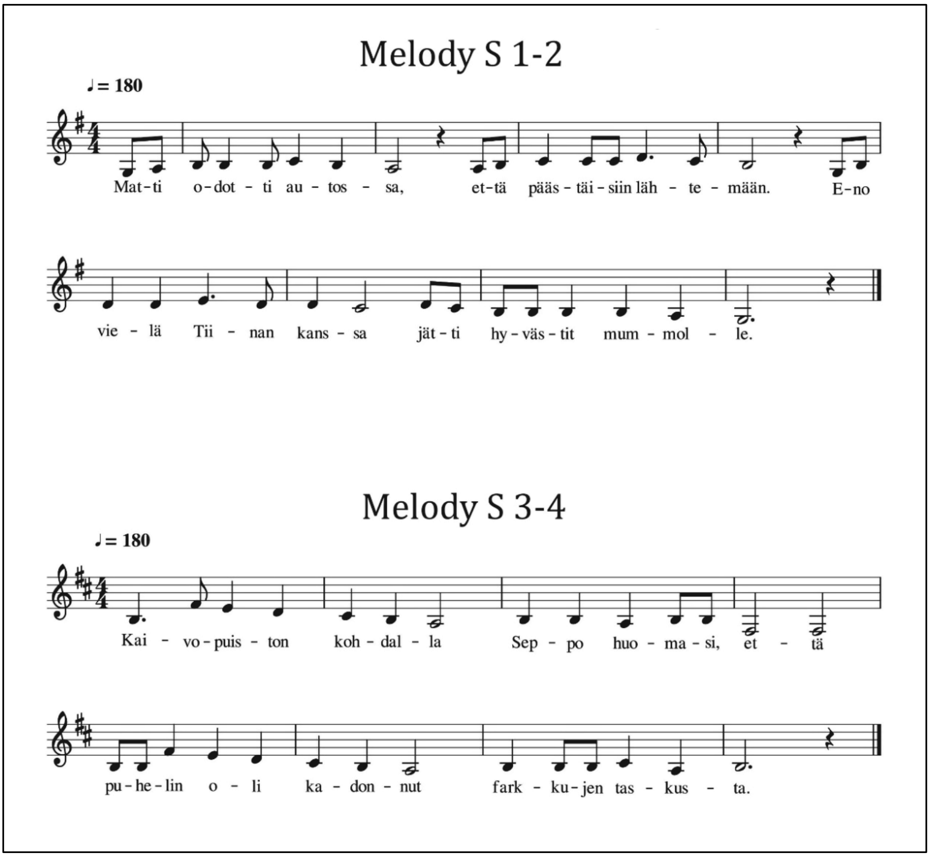

Figure 2 The notation of the melodies. S1-2 was used at the baseline and S3-4 at 6-month follow-up (reproduced with permission from the publisher, Leo et al., 2018).

On each trial, the task of the subject was to recall as much of the story as possible. To make the recall situation as natural and comfortable as possible in the sung condition, the subjects were given the option of recalling the story either by speaking or by singing. All recall performances were done by speaking as none of the subjects chose singing.

The scoring protocol for the SSSRT was similar as in the RBMT Story recall: two points for each correct word and one point for each partially correct or semantically similar word. The percentages of correct responses for each learning trial (T1, T2, T3) and the delayed (del) recall trial were calculated. For the serial position effect (SPE) analyses, these values were calculated separately for the first verse (V1), the middle verse (V3), and the last verse (V5), as well as for the primacy effect (PE, V1 minus V3) and the recency effect (RE, V5 minus V3).

For the chunking analyses, the item-level (individual words) scores were divided across the stories into chunks (defined as the number of consecutive words that were recalled correctly), and based on this, the average length of the chunks in the two tasks was then calculated.

3.3.3 Statistical analyses of behavioral data

In Studies I and II, the differences between the sung and spoken task performance were analyzed using mixed-model analyses of variance (ANOVA) as well as independent-samples and paired t-tests. In Study I, the mixed-model ANOVAs were performed for the difference between the sung and spoken task performance (sung minus spoken) for each learning trial with Time (acute / 6-month) as a within-subjects factor and Aphasia (non-aphasic / aphasic), Amusia (non-amusic / amusic), and Memory impairment (not memory-impaired / memory-impaired) as between-subjects factors. In Study II, the mixed-model ANOVAs were performed for the SPE analyses with Task (spoken / sung), Trial (T1 / T2 / T3 / del), and Verse (V1 / V3 / V5) as within-subject factors and Aphasia (non-aphasic / aphasic) as a between-subject factor and for the chunking analyses with Task (spoken / sung) and Trial (T1 / T2 / T3 / del) as within-subject factors and Aphasia (non-aphasic / aphasic) as a between-subject factor.

In Study III, demographic and clinical characteristics were analyzed with univariate ANOVAs and t-tests or non-parametric Kruskal-Wallis and Mann-Whitney tests, and chi-square tests. Longitudinal cognitive data were analyzed using mixed-model ANOVAs with Time as within-subject factor and Group and Aphasia as between-subject factors. Separate mixed-model ANOVAs were performed to determine (1) short-term (Time: acute / 3-month), (2) long-term (Time: acute / 6-month), (3) general [Group: MG (VMG and IMG combined) / ABG], and (4) specific (Group: VMG/IMG/ABG) effects of the intervention. Due to the emotional lability of the patients at the acute stage, the POMS data were analyzed cross-sectionally at each time point using univariate ANOVAs. In order to control for the potential effects of the two trial sites (Turku/ Helsinki) and those demographic and clinical variables, in which there were group differences, trial site, pre-stroke music listening, and cross-listening were included as covariates in the two-group (MG/ABG) and three-group (VMG/IMG/ABG) comparisons in the analyses of outcome measures. In addition, amusia was included as an additional covariate in the three –group comparisons. Post hoc tests of change scores (3-month minus acute, 6-month minus acute) were performed using the Bonferroni correction. All statistical analyses were performed using IBM SPSS Statistics 24. The level of statistical significance was set at $p < 0.05$. Missing values in the data were not replaced, but considered missing at random.

3.4 Magnetic resonance imaging

Structural MRI provides information on the size, shape, and integrity of brain structures, whereas functional MRI (fMRI) provides information about brain activity by detecting changes in blood flow. In Studies II and III, structural MRI was used verify the location and extent of the stroke lesion and to assess changes in grey matter and white matter volume using voxel-based morphometry and in the fine structure of WM tracts using diffusion tensor imaging and deterministic tractography.

3.4.1 MRI data acquisition (Studies II and III)

Structural MRI was acquired from altogether 75 patients at the acute and 6-month stages. Patients from Helsinki (N=33) were scanned using a 1.5 T Siemens Vision scanner (Siemens Medical Solutions, Erlangen, Germany) of the Department of Radiology in Helsinki University Central Hospital. Patients from Turku (N=42) were scanned using a 3 T Siemens Verio scanner (Siemens Medical Solutions, Erlangen, Germany) of the Medical Imaging Centre of Southwest Finland. High-resolution T1 images [Helsinki: flip angle = 15°, repetition time (TR) = 1900 ms, echo time = 3.68 ms, voxel size = 1.0 × 1.0 × 1.0 mm, Turku: flip angle = 9°, TR = 2300 ms, echo time = 2.98 ms, voxel size = 1.0 × 1.0 × 1.0 mm) were obtained and coupled with fluid-attenuated inversion recovery images that were used in localizing the lesion areas.

In Turku, also diffusion-weighted MRI scans (TR = 11,700 ms, echo time = 88 ms, acquisition matrix = 112×112, voxel size = 2.0 × 2.0 × 2.0 mm, 66 axial slices) with one non-diffusion weighted volume and 64 diffusion weighted volumes (b-values of 1000 s/mm2) were acquired for diffusion tensor imaging and single-shot T2*-weighted gradient-echo planar imaging sequence (flip angle = 80°, TR = 2010 ms, echo time = 30 ms, voxel size = 2.8 × 2.8 × 3.5 mm, slice thickness = 3.5 mm, 32 slices), with a total of 280 functional volumes were acquired for fMRI. The fMRI session included a 5-minute eyes-open resting-state condition and a passive listening task (task-fMRI) while auditory stimuli were presented through magnetic resonance-compatible headphones using Presentation software (Neurobehavioral Systems, version 16.3). In task-fMRI, a block design was used where the patients were presented 15-second excerpts of well-known Finnish songs with sung lyrics (Vocal, 6 blocks), without sung lyrics (Instrumental, 6 blocks), well-known Finnish poems (Speech, 6 blocks), and no auditory stimuli (Rest, 18 blocks). The order of the auditory blocks was randomized across subjects and time. The rest blocks were presented in between the auditory blocks.

3.4.2 MRI data preprocessing (Studies II and III)

MRI data were preprocessed using Statistical Parametric Mapping software (SPM8, Wellcome Department of Cognitive Neurology, UCL) under MATLAB 8.4.0 (The MathWorks Inc., Natick, MA, USA, version R2014b). The structural T1 images of each subject were reoriented to the anterior commissure and then processed using Unified Segmentation (Ashburner & Friston, 2005) with medium regularization. To better account for the lesions, cost function masking (Brett, Leff, Rorden & Ashburner ,2001) was applied to achieve accurate segmentation and optimal normalization of the lesioned grey matter and white matter tissue, with no post-registration lesion shrinkage or out-of-brain distortion (Ripollés et al., 2012). Using MRIcron (https://www.nitrc.org/projects/mricron, Rorden & Brett, 2000), cost function masking was performed by manually depicting the lesioned areas slice-by-slice to the T1 images of each subject. This approach has been well established in stroke patients (Crinion et al., 2007; Ripollés et al., 2012; Särkämö et al., 2014; Sihvonen et al., 2016). The segmented grey matter and white matter images were modulated to preserve the original signal strength and then normalized to the MNI space. After this, to reduce residual inter-individual variability, grey matter and white matter probability maps were smoothed using an isotropic spatial filter 6 mm. For fMRI data, the functional runs were first realigned and their mean image was calculated. The T1 image and its lesion mask were then co-registered to this mean functional image. The normalization parameters were again estimated using Unified Segmentation with cost function masking and were applied to the whole functional run to register it to MNI space. In this registration step, data were resampled into 2.0 × 2.0 × 2.0 mm voxel size. Finally, the normalized fMRI data was smoothed using an 8 mm isotropic spatial filter kernel.

3.4.3 Deterministic tractography (Study II)

In Study II, deterministic tractography was used to evaluate the relationship between task performance and white matter tracts. First, corrections for eddy current distortions and head motion were carried out using FMRIB Software Library (University of Oxford, FSL v5.0.8, www.fmrib.ox.ac.uk/fsl, Smith et al., 2004). Second, the gradient matrix was rotated using FMRIB Software Library's fdt rotate bvecs to provide more accurate estimate of diffusion tensor orientations (Leemans & Jones, 2009). Then, the Brain Extraction Tool was used to perform the brain extraction (Smith, 2002) and diffusion tensors were reconstructed using the linear least-squares algorithm included in Diffusion Toolkit 0.6.2.2 (Wang, Benner, Sorensen & Wedeen, 2007; trackvis.org/dtk, Martinos Center for Biomedical Imaging, Massachusetts General Hospital). Finally, fractional anisotropy and mean and radial diffusivity maps for each patient were calculated using the eigenvalues extracted from the diffusion tensors.

Based on previous studies linking them to verbal learning or verbal memory (Chiou, Genova, & Chiaravalloti, 2016; López-Barroso et al., 2013; Mabbott, Rovet, Noseworthy, Smith, & Rockel, 2009; McDonald et al., 2008; Reggente et al., 2018), the following four frontotemporal white matter tracts in both hemispheres were dissected using TrackVis (version 0.6.0.1, Build 2015.04.07) and included in the deterministic tractography analyses: arcuate fasciculus (AF), inferior fronto-occipital fasciculus (IFOF), inferior longitudinal fasciculus (ILF), and uncinate fasciculus (UF). After dissection, statistical information on tract volume and fractional anisotropy, which is a scalar value between zero and one that provides an index of structural integrity of the tract, of each white matter tract was collected using a MATLAB toolbox, "along-tract statistics" (Colby et al., 2012). The tract volume and fractional anisotropy values were imported to IBM SPSS Statistics 24. These values were then analyzed to evaluate the relationship between the white matter tract parameters and behavioral performance (PE, RE, and chunking in the SSSRT) using two-tailed Pearson correlation analysis in aphasic and non-aphasic patients. Standard false discovery rate (FDR) correction was applied to control for multiple correlations.

3.4.4 Voxel based morphometry (Studies II and III)

Voxel-based morphometry is a computational structural MRI analysis approach that allows estimation of grey matter and white matter differences in specific brain regions between groups and over time, through a voxel-wise comparison of multiple brain images (Ashburner & Friston, 2000). The voxel-based morphometry analysis was performed using SPM8 under MATLAB 8.4.0. The preprocessed and modulated grey matter and white matter images were entered into a second-level analysis, which was different in Study II and in Study III.

In Study II, t-tests were used to assess the relationship between the behavioral performance (PE, RE, and chunking in the SSSRT) and the grey matter volume across the entire grey matter space within the aphasic group and in the aphasic vs. non-aphasic group. Age, gender, and total intracranial volume were added as nuisance covariates (Barnes et al., 2010). All results were thresholded at a whole-brain uncorrected $p < 0.001$ threshold with a cluster extent of >100 contiguous voxels (Lieberman & Cunningham, 2009). To evaluate which grey matter correlates were facilitating the behavioral performance, partial correlations with two-tailed false discovery rate (FDR) corrected p-values controlling for age, sex and total intracranial volume were calculated for each significant cluster separately for aphasic and non-aphasic patients.

In Study III, a flexible factorial analysis was performed with Time (acute / 6-month) and Group, with either three (VMG / IMG / ABG) or two (MG / ABG) levels, as factors, and scanner type, age, sex, and total intracranial volume as additional covariates.

Thus, altogether four Group (VMG > ABG, IMG > ABG, VMG > IMG, MG > ABG) x Time (6-month > acute) interactions were calculated. Separate analyses were also performed within the aphasic and non-aphasic patients. All results were thresholded at an uncorrected $p < 0.005$ threshold with a minimal cluster size set to 100 voxels. Only clusters surviving a family-wise error (FWE) correction at $p < 0.05$ at cluster level are reported.

In the second-level results, the Automated Anatomical Labeling Atlas (Tzourio-Mazoyer et al., 2002) was used to identify neuroanatomical areas provided within the xjView toolbox (http://www.alivelearn.net/xjview/).

3.4.5 fMRI functional connectivity analyses (Study III)

In Study III, functional connectivity analyses of the fMRI data were performed with group-level spatial independent component analysis using the Group independent component analysis of fMRI Toolbox (GIFT) software (http://mialab.mrn.org/software/gift/). The spatial components of independent component analysis were extracted from the resting state -fMRI and task-fMRI runs. After performing intensity normalization of the preprocessed fMRI images, data were concatenated and, following previous studies (Smith et al., 2009; López-Barroso et al., 2015), reduced to 20 temporal dimensions using principal component analysis (PCA) and then analyzed using the infomax algorithm (Bell & Sejnowski, 1995). The default mode network was identified from the spatial components of independent component analysis representing the different networks, and selected for further analyses based on the pattern of the voxel-based morphometry results (see Results section). To obtain whole brain group-wise statistics in the resting state-fMRI, the spatial maps of the default mode network from all patients were submitted to a second-level flexible factorial analyses with Time (acute / 6-month) and Group (VMG / IMG / ABG) as factors. In the task-fMRI, the time course of the default mode network was fitted to a statistical parametric mapping (SPM) model that included the Vocal, Instrumental, and Speech conditions as regressors. The beta values yielded representing the engagement of default mode network during each condition, which were then analyzed with SPSS using mixed-model ANOVAs with Time (acute / 6-month) and Group (VMG / IMG / ABG) as factors. Scores from Boston diagnostic aphasia examination severity rating scale and Barcelona music reward questionnaire were included as additional covariates. Statistical maps were thresholded at a voxel-level uncorrected $p<0.005$ threshold with k-extent ≥100 and a familywise error rate (FWE) corrected $p<0.05$ at the cluster level, a combination which has been shown to produce a desirable balance between types I and II error rates, comparable to false discovery rate (FDR) (Lieberman & Cunningham, 2009). To determine the link between the behavioral outcome and the voxel-based morphometry -functional connectivity results, correlation analyses (Pearson, two-tailed, FDR-corrected) were performed between changes (3-month minus acute, 6-month minus acute) in language skills and verbal memory and the clusters showing functional connectivity or volume changes between the groups.

4 RESULTS

4.1 Patient characteristics

In Studies I and II, based on their performance on the language tests and on the Boston diagnostic aphasia examination severity rating scale scores, 14 patients (45%) were classified as aphasic and 17 as non-aphasic. In the aphasic subgroup, all of the patients had a left hemisphere lesion and the severity of their language impairment was primarily mild (10 subjects received from Boston diagnostic aphasia examination severity rating scale score 4 and 4 subjects received score 3). As shown in Table 2, the aphasic and non-aphasic groups did not differ in age, gender, formal musical training, pre-stroke musical hobbies, education, lesion size, NIHSS scores, or RBMT-Story recall performance, but the aphasic group had lower scores in Auditory-Verbal Learning Task learning [$t(29) = 2.9$, $p = 0.006$] and delayed recall [$t(29) = 2.1$, $p = 0.049$] at the acute stage, as expected.

In Study I, the patients were also classified based on amusia and memory impairment. Based on their performance on the MBEA Scale and Rhythm subtests, 16 patients (52%) were classified amusic and 15 as non-amusic. Based on their performance on the verbal memory tests (RBMT Story recall and Auditory-verbal learning task), 16 patients (52%) were classified as having more severe memory impairment (referred to hereafter as memory-impaired) and 15 as having less severe or no memory impairment (referred to hereafter as not memory-impaired). The amusic and non-amusic groups and the memory-impaired and not memory-impaired groups were quite well matched on demographic and clinical variables (Table 2). The only significant differences were that at the acute stage the amusic group higher NIHSS scores than the non-amusic group ($p = 0.018$) and the memory-impaired group had lower frequency of pre-stroke music listening than the not memory-impaired group ($p = 0.046$).

In Study III, there were no statistically significant differences between MG and ABG or between VMG, IMG, and ABG at the acute stage baseline in most demographic and clinical characteristics and pre-stroke leisure activities (Table 3). Pre-stroke music listening frequency showed a difference in two-group (Mann-Whitney $U = 533.50$, $p = 0.003$; MG > ABG) and three-group [Kruskal-Wallis $H = 11.81$, $p = 0.003$] comparisons, with more pre-stroke music listening in IMG than in VMG ($p = 0.007$) or ABG ($p = 0.010$). In addition, the proportion of amusic patients differed between the three groups [$\chi^2(2) = 9.29$, $p = 0.010$], with less amusics in IMG than in VMG ($p = 0.034$) or ABG ($p = 0.001$). The groups were also comparable in in the amount of received motor, speech or cognitive rehabilitation at the 3-month and 6-month stage. The frequency of music and audiobook listening differed highly significantly between the groups, both during the intervention and follow-up periods. Listening frequency was higher for music in VMG and IMG than in ABG, and for audiobooks in ABG than in VMG and IMG at the 3-month stage ($p < 0.001$ in all) and, to a lesser extent, at the 6-month stage ($p < 0.092$ in all).

Table 2. Characteristics of the patients in Studies I and II (reproduced with permission from the publishers, Leo et al., 2018, 2019)

	Aphasic N = 14	**Non-aphasic** N = 17	**p-value**	**Amusic** N = 16	**Non-amusic** N = 15	**p-value**	**Memory-impaired** N= 16	**Not memory-impaired** N=15	**p-value**
Demographical variables									
Age (years)	51.4 (17.7)	54.4 (11.3)	0.564 (t)	55.4 (12.9)	50.5 (15.6)	0.365 (t)	54.4 (11.3)	51.6 (17.0)	0.599 (t)
Gender (male/female)	9 / 5	10 / 7	0.756 (χ²)	12 / 4	7 / 8	0.106 (χ²)	9 / 7	10 / 5	0.552 (χ²)
Education (years)	15.1 (4.3)	14.1 (2.4)	0.428 (t)	14.0 (3.4)	15.1 (3.3)	0.344 (t)	14.6 (4.1)	14.5 (2.6)	0.375 (t)
Pre-stroke musical background									
Formal music training (yes/no)	4 / 10	4 / 13	0.750 (χ²)	4 / 12	4 / 11	0.916 (χ²)	3 / 13	5 /10	0.354 (χ²)
Active singing or playing (yes/no)	7 / 7	8 / 9	0.870 (χ²)	6 / 10	9 / 6	0.210 (χ²)	8 / 8	7 / 8	0.853 (χ²)
BMRQ score (max.100)	77.6 (11.5)	74.0 (13.7)	0.435 (t)	77.0 (10.7)	74.2 (14.7)	0.547 (t)	74.8 (13.2)	76.5 (12.4)	0.711 (t)
Pre-stroke leisure activities									
Music listening frequency[a]	4.4 (1.3)	4.8 (0.8)	0.245 (t)	4.3 (1.3)	4.9 (0.4)	0.106 (t)	4.3 (1.2)	4.9 (0.3)	0.046 (t)
Radio listening frequency[a]	2.4 (1.3)	3.0 (1.8)	0.328 (t)	2.9 (1.7)	2.5 (1.6)	0.489 (t)	2.3 (1.6)	3.3 (1.4)	0.075 (t)
Reading frequency[a]	3.9 (1.4)	3.8 (2.0)	0.958 (t)	3.9 (1.7)	3.7 (1.8)	0.747 (t)	3.4 (1.9)	4.3 (1.4)	0.183 (t)
Clinical variables (acute post-stroke)									
Lesion laterality (left/right)	14 / 0	6 / 11	0.001 (χ²)	8 / 8	12 / 3	0.081 (χ²)	12 / 4	8 / 7	0.208 (χ²)
Lesion size (cm³)	39.0 (50.4)	65.3 (56.2)	0.186 (t)	63.6 (58.1)	42.6 (50.0)	0.292 (t)	60.9 (55.5)	45.5 (54.0)	0.439 (t)
Stroke type (infarct / hemorrhage)	9 / 5	13 / 4	0.457 (χ²)	12 / 4	10 / 5	0.609 (χ²)	10 / 6	12 / 3	0.283 (χ²)
NIHSS score (max. 42)	3.9 (2.3)	5.5 (3.6)	0.158 (t)	6.0 (3.6)	3.4 (1.9)	0.018 (t)	5.0 (3.2)	4.5 (3.2)	0.644 (t)
BDAE Aphasia Severity Rating Scale	3.7 (0.5)	4.8 (0.4)	< 0.001(t)	4.0 (0.7)	4.6 (0.5)	0.013 (t)	4.1 (0.7)	4.5 (0.6)	0.057 (t)
MBEA Scale and Rhythm avg (%)	73.9 (9.9)	72.1 (17.1)	0.720 (t)	61.3 (8.9)	85.3 (4.5)	< 0.001 (t)	71.9 (14.4)	74.0 (14.1)	0.682 (t)
AVLT Learning score (3 trials, max. 30)	15.6 (4.0)	20.0 (4.2)	0.006 (t)	17.4 (5.4)	18.7 (3.7)	0.466 (t)	15.4 (4.0)	20.9 (3.4)	< 0.001 (t)
AVLT Delayed recall score (max. 10)	3.4 (2.4)	5.4 (2.7)	0.049 (t)	4.1 (2.9)	4.9 (2.6)	0.459 (t)	2.9 (2.5)	6.1 (1.9)	< 0.001 (t)
RBMT Story recall immediate (max. 42)	12.0 (6.0)	15.5 (7.6)	0.174 (t)	14.1 (7.4)	13.7 (6.8)	0.899 (t)	8.8 (3.5)	19.3 (5.6)	< 0.001 (t)
RBMT Story recall delayed (max. 42)	8.6 (5.3)	13.1 (8.6)	0.104 (t)	10.1 (7.7)	12.1 (7.4)	0.482 (t)	5.8 (3.6)	16.7 (6.5)	< 0.001 (t)

Data are reported as mean (SD) unless otherwise stated. Abbreviations: t = independent-samples t-test, χ² = chi-square test, AVLT = Auditory-Verbal Learning Task, BDAE= Boston Diagnostic Aphasia Examination, BMRQ = Barcelona Music Reward Questionnaire, MBEA= Montreal Battery of Evaluation of Amusia, NIHSS = National Institute of Health Stroke Scale, RBMT = Rivermead Behavioural Memory Test

[a]Likert scale 1-7 (1 = not at all, 7 = daily); [b]Likert scale 0-5 (0 = no usable speech or comprehension, 5 = minimal speech handicaps)

Table 3. Baseline characteristics, listening activity, and received rehabilitation of the patients in Study III (Sihvonen & Leo et al., in revision)

	MG N = 50	VMG N = 27	IMG N = 23	ABG N = 33	MG / ABG p-value	MG / IMG / ABG p-value
Demographic characteristics						
Age (years)	55.7 (12.0)	54.9 (13.4)	56.7 (10.3)	59.8 (11.6)	0.123 (t)	0.267 (F)
Gender (male / female)	30 / 20	15 / 12	15 / 8	16 / 17	0.302 (χ^2)	0.464 (χ^2)
Education (years)	13.1 (4.2)	12.8 (4.5)	13.5 (3.8)	12.1 (3.5)	0.229 (t)	0.405 (F)
Pre-stroke leisure activities[a]						
Music listening	4.4 (1.1)	4.2 (1.2)	4.7 (0.9)	3.6 (1.5)	0.003 (U)	0.003 (H)
Radio listening	3.4 (1.7)	3.3 (1.9)	3.4 (1.4)	3.8 (1.4)	0.280 (U)	0.545 (H)
Reading	4.1 (1.6)	3.9 (1.7)	4.3 (1.5)	4.6 (0.8)	0.159 (U)	0.138 (H)
Clinical characteristics						
Time from stroke to baseline (days)	6.6 (3.9)	6.1 (2.6)	7.2 (5.0)	7.3 (4.1)	0.481 (t)	0.502 (F)
Lesion laterality (left / right)	25 / 25	15 / 12	10 / 13	16 / 17	0.893 (χ^2)	0.690 (χ^2)
Lesion size (cm^3)	48.9 (50.3)	46.8 (50.8)	51.4 (50.9)	45.1 (45.5)	0.725 (t)	0.893 (F)
Aphasia (yes / no)[b]	15 / 35	10 / 17	5 / 18	14 / 19	0.245 (χ^2)	0.269 (χ^2)
Aphasia severity[b]	3.3 (1.2)	3.1 (1.3)	3.6 (0.9)	3.2 (1.1)	0.561 (U)	0.593 (H)
Amusia (yes / no)[c]	27 / 22	19 / 7	8 / 15	23 / 10	0.184 (χ^2)	0.010 (χ^2)
Listening and rehabilitation (3-month)						
Listening to music[d]	5.0 (0.0)	5.0 (0.0)	5.0 (0.0)	2.7 (2.1)	<0.001 (U)	<0.001 (H)
Listening to audio books[d]	0.1 (0.5)	0.0 (0.2)	0.1 (0.7)	4.6 (1.0)	<0.001 (U)	<0.001 (H)
Cross-listening of material[d]	0.1 (0.5)	0.0 (0.2)	0.1 (0.7)	2.7 (2.1)	<0.001 (U)	<0.001 (H)
Received motor rehabilitation[e]	22.7 (31.9)	28.6 (34.3)	15.7 (28.0)	15.5 (21.5)	0.609 (U)	0.180 (H)
Received speech rehabilitation[e]	3.9 (6.8)	4.8 (7.8)	2.9 (5.4)	3.8 (7.0)	0.655 (U)	0.804 (H)
Received cognitive rehabilitation[e]	3.8 (9.8)	5.6 (12.8)	1.7 (3.7)	1.5 (3.1)	0.464 (U)	0.342 (H)
Listening and rehabilitation (6-month)						
Listening to music[d]	4.0 (1.4)	4.2 (1.2)	3.7 (1.5)	2.9 (1.6)	0.004 (U)	0.008 (H)
Listening to audio books[d]	0.3 (0.9)	0.2 (0.8)	0.4 (1.1)	2.8 (2.0)	<0.001 (U)	<0.001 (H)
Received motor rehabilitation[e]	32.7 (44.1)	39.6 (44.8)	24.8 (42.8)	27.5 (39.6)	0.723 (U)	0.282 (H)
Received speech rehabilitation[e]	5.3 (11.3)	6.2 (12.7)	4.2 (9.7)	4.6 (7.5)	0.482 (U)	0.764 (H)
Received cognitive rehabilitation[e]	5.1 (10.5)	7.1 (13.4)	2.9 (5.3)	4.4 (7.4)	0.791 (U)	0.559 (H)

Data are mean (SD) unless otherwise stated. Significant group differences are shown in bold. Abbreviations: ABG = Audio book group, F = one-way ANOVA, H = Kruskal-Wallis test, IMG = Instrumental music group, MG = Music group (VMG & IMG combined), t = t-test, U = Mann-Whitney test, VMG = Vocal music group, χ^2 = chi-square test
[a]Likert scale 0-5 (0= never, 1= rarely, 2= once a month, 3= once a week, 4= 2-3 times a week, 5= daily); [b]Classification based on Boston diagnostic aphasia examination severity rating scale: scores 0-4 = aphasia, score 5 = no aphasia. For aphasic patients, the mean score is shown; [c]Classification based on the MBEA Scale & Rhythm subtest average score (<75% cut-off); [d]Likert scale 0-5 (0= never, 1= rarely, 2= once a month, 3= once a week, 4= 2-3 times a week, 5= daily); [e]Number of therapy sessions (motor: physical or occupational therapy, speech: speech therapy, cognitive: neuropsychological rehabilitation)

There were no significant differences in music or audiobook listening between VMG and IMG. Even though at group level the listening frequencies followed the study protocol, there was a significant difference between the groups in cross-listening

(listening to material not part of the protocol: music in ABG, audiobooks in VMG and IMG) during the intervention period ($H = 42.64$, $p < 0.001$), with ABG showing more cross-listening than VMG and IMG ($p < 0.001$ in both). Thus, pre-stroke music listening and cross listening were included as covariates in the two-group (MG / ABG) and three-group (VMG / IMG / ABG) comparisons and amusia as an additional covariate in the three-group comparisons.

4.2 Effect of sung presentation on verbal learning and recall (Study I)

In the SSSRT, a direct comparison between the sung and spoken versions using paired t-tests did not show any significant differences in performance between the two tasks across all subjects either in the 1st (T1), 2nd (T2), or 3rd (T3) learning trial or in delayed recall at the acute post-stroke stage (see Figure 3A). Separate paired t-tests within the aphasic, amusic, and memory-impaired subgroups did not yield any significant effects at the acute stage. At the 6-month post-stroke stage (see Figure 3B), performance on the sung task was better than on the spoken task across all subjects in the 2nd [$t(30) = -2.8$, $p = 0.009$] and 3rd [$t(29) = -2.4$, $p = 0.023$] learning trials and in the delayed recall [$t(30) = -3.6$, $p = 0.001$]. Separate paired t-tests in the aphasic, amusic, and memory-impaired subgroups indicated that performance on the sung task was better than on the spoken task in the 2nd [$t\ (13) = -2.2$, $p = 0.045$] and 3rd [$t(13) = -2.8$, $p = 0.015$] learning trials in the aphasic subgroup, in the delayed recall trial [$t(15) = -2.4$, $p = 0.029$] in the amusic subgroup, and in the 2nd learning [$t\ (15) = -2.2$, $p = 0.048$] and delayed [$t(15) = -2.5$, $p = 0.022$] trials in the memory-impaired subgroup.

As shown in Figure 4, mixed-model ANOVAs on the difference between the sung and spoken task performance (sung minus spoken) yielded a significant Time x Aphasia interaction for all the learning trials [1st: $F(1,23) = 5.5$, $p = 0.028$; 2nd: $F(1,23) = 12.7$, $p = 0.002$; 3rd: $F(1,22) = 5.2$, $p = 0.033$]. Aphasic patients showed a larger increase in the sung over spoken effect from the acute to the 6-month stage compared to non-aphasic patients. A significant Time x Amusia interaction was found for the delayed recall trial [$F(1,23) = 5.7$, $p = 0.025$], with amusic patients showing a larger increase in the sung over spoken effect from the acute to the 6-month stage compared to non-amusic patients. Also, a significant three-way Time x Aphasia x Memory impairment interaction was found for the delayed recall trial [$F\ (1,23) = 5.2$, $p = 0.032$]. Separate mixed-model ANOVAs within the not memory-impaired and memory-impaired and subgroups showed that the longitudinal sung over spoken delayed recall effect in aphasic vs. non-aphasic patients was significant only within the not memory-impaired subgroup [Time x Aphasia: $F(1,13) = 5.6$, $p = 0.034$]. This result suggests that those aphasic patients who did not have concurrent memory impairment benefited most from the sung presentation of the material in delayed recall during the follow-up.

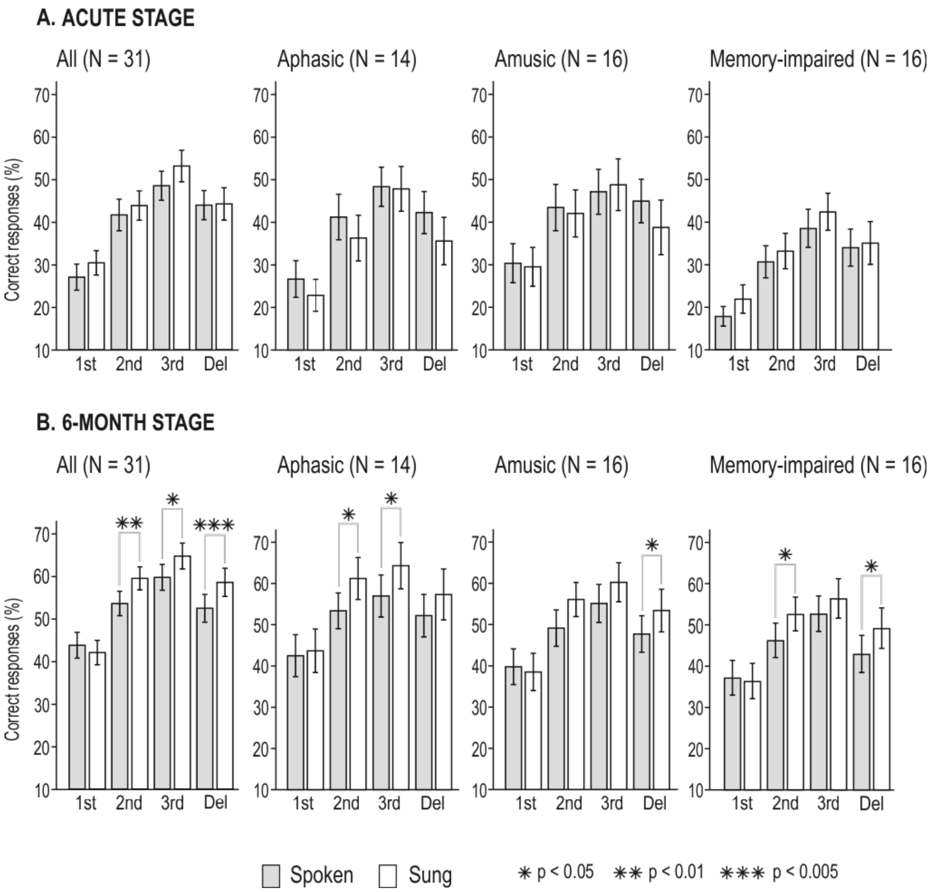

Figure 3 Percentage of correct responses (mean ± SEM) of the patients on the SSSRT separately for the spoken (grey bars) and sung (white bars) conditions at the acute (panel A) and 6-month (panel B) post-stroke stage. Results are shown for all patients (upper panel) and for aphasic and amusic subgroups (lower panel). Significant results in paired t-tests are indicated with asterisks (reproduced with permission from the publisher, Leo et al., 2018).

In summary, the results of Study I showed that stroke patients can benefit from the sung melody as a mnemonic aid in recalling novel stories, especially starting from the 2nd learning trial, but only at later (6-month) recovery stage. Moreover, the benefit of the sung melody on verbal learning was seen especially in patients with mild aphasia, in whom this effect increased greatly from the acute to the 6-month stage.

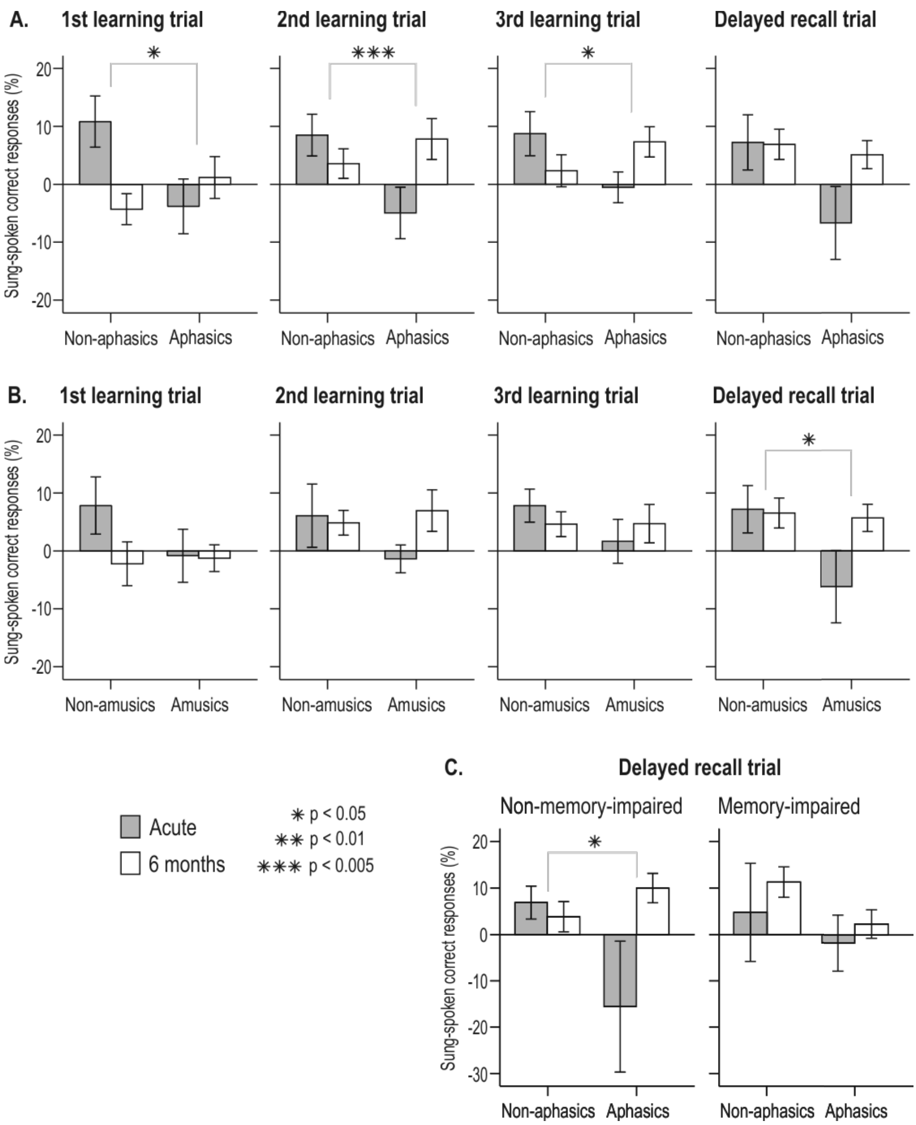

Figure 4 Difference between the percentage of correct responses in the sung and spoken parts (sung minus spoken) of the SSSRT at acute (grey bars) and 6-month (white bars) post-stroke stage. Results are shown (mean ± SEM) for groups no aphasia / aphasia (upper panel) and no amusia / amusia and groups of no aphasia / aphasia (lower panel). Significant Time x Group interactions in mixed-model ANOVA are shown with asterisks (reproduced with permission from the publisher, Leo et al., 2018).

4.3 Cognitive and neural mechanisms underlying of benefit of sung presentation on verbal learning and recall (Study II)

4.3.1 Serial position effect (SPE) results

The verse-level SSSRT performance of the patients at the 6-month stage is shown in Figure 5 and Table 4. A mixed-model ANOVA of the 6-month SSSRT data with Task (spoken / sung), Trial (T1 / T2 / T3 / del), and Verse (V1 / V3 / V5) as within-subject factors and Aphasia (non-aphasic / aphasic) as a between-subject factor yielded a significant Verse main effect [$F(2, 56) = 17.2$, $p < 0.001$], indicating better recall of V1 than V3 and of V5 than V3, which points to a general primacy effect (PE) and recency effect (RE) in the SSSRT. There was also a significant Trial x Verse interaction [$F(3.6, 99.5) = 4.4$, $p = 0.004$], indicating more uniform recall of the verses in the delayed recall trial than in the learning trials, and a Task x Verse x Aphasia interaction [$F(1.5, 40.9) = 3.7$, $p = 0.046$].

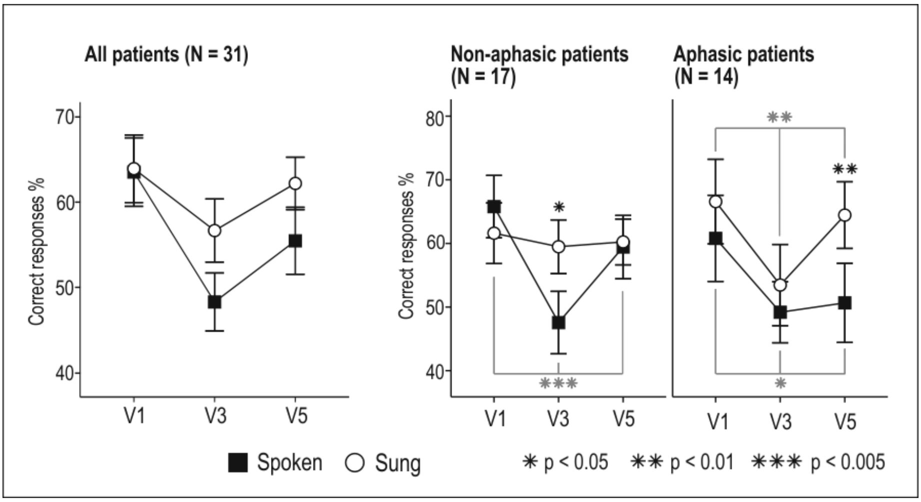

Figure 5 Percentage of correct responses (mean ± SEM) in the first (V1), middle (V3) and last (V5) verses of the sung (white circles) and spoken (black squares) story recall tasks. Data are shown across all patients (left) and within the groups without aphasia and aphasia (right). Grey asterisks denote significant Verse main effects in mixed-model ANOVAs in the sung and spoken the tasks. Black asterisks denote significant differences between the sung and spoken tasks for individual verses in paired t-tests (reproduced with permission from the publisher, Leo et al., 2019).

Table 4. Verse-level SSSRT performance at the 6-month stage in Study II (reproduced with permission from the publisher, Leo et al., 2019).

Task	Trial	Verse	All patients N = 31	Aphasic N = 14	Non-aphasic N = 17
Spoken	1st learning trial (T1)	V1	57.7 (23.0)	54.8 (20.4)	60.2 (25.2)
		V3	35.2 (27.7)	39.3 (31.1)	31.8 (25.0)
		V5	44.4 (28.5)	38.3 (32.2)	49.5 (25.0)
	2nd learning trial (T2)	V1	67.6 (24.9)	64.7 (28.6)	70.0 (22.1)
		V3	45.7 (23.7)	45.4 (22.9)	45.8 (25.0)
		V5	61.9 (23.1)	58.8 (23.1)	64.4 (23.4)
	3rd learning trial (T3)	V1	72.3 (25.1)	65.4 (31.4)	78.1 (17.4)
		V3	58.1 (22.4)	56.3 (21.1)	59.5 (24.0)
		V5	62.3 (28.0)	58.1 (29.1)	65.8 (27.5)
	Delayed recall trial (del)	V1	56.5 (27.0)	58.3 (31.1)	54.9 (23.9)
		V3	54.2 (21.4)	55.6 (21.0)	53.0 (22.3)
		V5	53.2 (25.6)	47.4 (30.1)	58.0 (20.9)
	Average across trials	V1	63.5 (22.4)	60.8 (25.3)	65.8 (20.2)
		V3	48.3 (19.0)	49.2 (18.0)	47.5 (20.3)
		V5	55.5 (21.9)	50.6 (23.3)	59.4 (20.5)
	PE (V1 minus V3)		15.2 (20.3)	11.6 (14.8)	18.2 (23.9)
	RE (V5 minus V3)		7.2 (16.6)	1.5 (14.7)	11.9 (17.0)
Sung	1st learning trial (T1)	V1	49.3 (26.5)	51.4 (33.0)	47.8 (20.6)
		V3	34.9 (24.8)	33.7 (21.5)	35.9 (27.8)
		V5	50.3 (22.1)	51.0 (25.0)	49.7 (20.2)
	2nd learning trial (T2)	V1	69.8 (22.4)	76.9 (20.2)	64.0 (23.0)
		V3	60.9 (24.2)	60.9 (31.3)	60.8 (17.4)
		V5	65.2 (19.0)	67.2 (21.3)	63.6 (17.5)
	3rd learning trial (T3)	V1	71.4 (24.8)	73.5 (27.6)	69.5 (22.7)
		V3	66.6 (23.9)	64.5 (26.4)	68.3 (22.1)
		V5	68.3 (19.5)	73.1 (21.5)	64.2 (17.2)
	Delayed recall trial (del)	V1	62.5 (25.3)	64.5 (29.6)	60.9 (22.0)
		V3	62.5 (26.1)	54.6 (31.7)	69.0 (19.0)
		V5	63.9 (18.5)	66.6 (20.3)	61.8 (17.2)
	Average across trials	V1	63.9 (21.7)	66.6 (24.9)	61.6 (19.1)
		V3	56.7 (20.3)	53.5 (23.9)	59.5 (16.9)
		V5	62.2 (16.9)	64.4 (19.6)	60.2 (14.4)
	PE (V1 minus V3)		7.3 (19.7)	13.1 (15.6)	2.1 (21.9)
	RE (V5 minus V3)		5.5 (17.2)	11.0 (15.4)	0.8 (17.6)

Data are mean (SD). PE = primacy effect, RE = recency effect, V1 = 1st verse, V3 = 3rd (middle) verse, V5 = 5th (last) verse

This latter three-way interaction was followed up by separate ANOVAs performed in the non-aphasic and aphasic patients for the spoken and sung tasks. In the non-aphasic patients, the Verse effect was significant in the spoken task [$F(2, 32) = 7.9$, $p = 0.002$], with post hoc tests showing better recall of V1 than V3 (PE; $p = 0.006$) and of V5 than V3 (RE; $p = 0.011$), but not in the sung task. In the aphasic patients, the Verse effect was significant in both the spoken task [$F(2, 26) = 5.5$, $p = 0.010$] and the sung task [$F(2, 26) = 6.3$, $p = 0.006$]. Post hoc tests showed that the aphasic patients recalled V1 better than V3 (PE; $p = 0.012$) and also V1 better than V5 ($p = 0.014$) in the spoken task, whereas they recalled V1 better than V3 (PE; $p = 0.008$) and V5 better than V3 (RE; $p = 0.019$) in the sung task.

The differences in the SPE patterns of the tasks were further analyzed first with paired t-tests comparing the sung and spoken tasks for each verse (V1 / V3 / V5) and for the PE (V1 minus V3) and RE (V5 minus V3) within the non-aphasic and aphasic groups. Additionally, independent-samples t-tests were used to compare the non-aphasic and aphasic groups on the difference between the tasks (sung minus spoken) for each verse (V1 / V3 / V5), PE, and RE. The paired t-tests showed that in the non-aphasic group, recall was better in the sung than spoken task for V3 [$t(15) = 2.7$, $p = 0.018$], and both the PE and the RE were marginally smaller in the sung than spoken task [$t(15) = 2.0$, $p = 0.063$ and $t(15) = 2.1$, $p = 0.055$, respectively]. In contrast, the aphasic group showed better recall of V5 in the sung than spoken task [$t(13) = 3.3$, $p = 0.006$]. The independent-samples t-tests showed that compared to the non-aphasic group, the aphasic group had better recall of V5 [$t(28) = 2.5$, $p = 0.019$] and also larger RE [$t(28) = 2.6$, $p = 0.014$] in the sung than spoken task.

In summary, these results suggest that in non-aphasic patients the sung presentation leads to more stable recall across the story, reducing the SPE and facilitating the recall of the middle part compared to the spoken presentation. In contrast, in aphasic patients, the sung presentation enhances the recall of the last part of the story, resulting in larger RE, compared to the spoken presentation.

4.3.2 Chunking results

The average chunk lengths in the SSSRT are presented in Figure 6 and Table 5. A mixed-model ANOVA with Task (spoken / sung) and Trial (T1 / T2 / T3 / del) as within-subject factors and Aphasia (non-aphasic / aphasic) as a between-subject factor showed a significant Trial main effect [$F(1.8, 53.1) = 15.8$, $p < 0.001$], indicating progressively increasing chunk length during the learning trials, as well as a significant Task x Aphasia interaction [$F(1, 29) = 7.4$, $p = 0.011$]. The Task x Aphasia interaction was further explored first with paired t-tests comparing the chunk length between the sung and spoken tasks within the non-aphasic and aphasic groups and then with an independent-samples t-test comparing the non-aphasic and aphasic groups on the chunk length difference between the tasks (sung minus spoken). The paired t-tests showed an opposite, marginally significant pattern of recall effects in the two tasks, with the non-aphasic group recalling longer chunks in the sung than spoken task [$t(13) = 2.0$, $p = 0.073$] and the aphasic group recalling longer chunks in the spoken than sung task [$t(15) = 1.8$, $p = 0.097$]. The independent-samples t-test showed that the aphasic group recalled longer chunks in the sung than spoken task compared to the non-aphasic group [$t(28) = 2.7$, $p = 0.013$]. Together, these results suggest that in aphasic patients sung presentation leads to more effective chunking of verbal material.

Table 5. Chunk lengths in the SSSRT at the 6-month stage in Study II (reproduced with permission from the publisher, Leo et al., 2019).

Task	Trial	All patients N = 31	Aphasic N = 14	Non-aphasic N = 17
Spoken	1st learning trial (T1)	3.1 (1.5)	3.1 (1.6)	3.1 (1.4)
	2nd learning trial (T2)	3.7 (1.6)	3.5 (1.8)	3.8 (1.5)
	3rd learning trial (T3)	4.9 (3.0)	4.3 (2.1)	5.5 (3.5)
	Delayed recall trial (del)	3.8 (2.5)	3.3 (1.9)	4.2 (2.8)
	Average across trials	3.9 (1.9)	3.6 (1.6)	4.2 (2.1)
Sung	1st learning trial (T1)	2.8 (1.2)	3.1 (1.5)	2.5 (0.8)
	2nd learning trial (T2)	3.3 (1.6)	3.5 (2.1)	3.1 (0.9)
	3rd learning trial (T3)	5.6 (4.5)	6.5 (6.2)	4.7 (2.0)
	Delayed recall trial (del)	4.2 (2.7)	4.7 (3.6)	3.8 (1.5)
	Average across trials	4.0 (2.1)	4.4 (2.8)	3.6 (1.1)

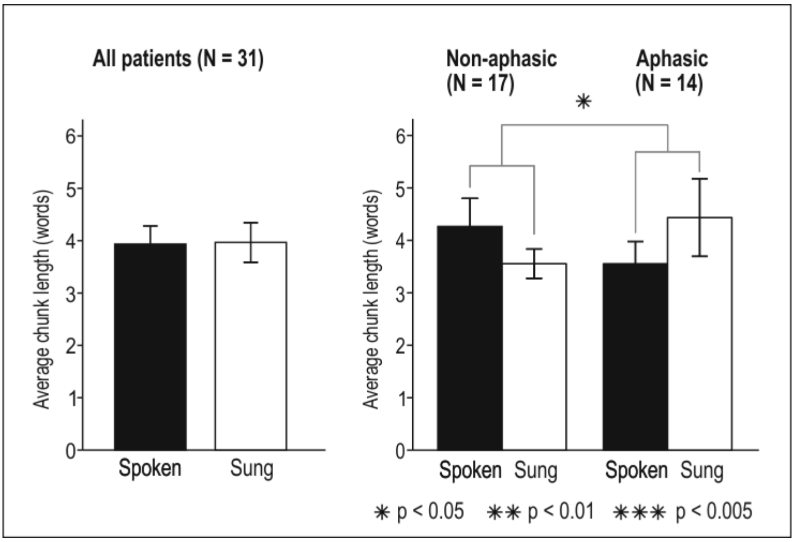

Figure 6 Average length of recalled chunks at the 6-month stage in Study II (mean ± SEM) in the spoken (black) and sung (white) tasks in all patients (left) and in non-aphasic patients and aphasic patients (right). Significant Task x Group interaction indicated with an asterisk (reproduced with permission from the publisher, Leo et al., 2018).

4.3.3 Neural correlates of the serial position and chunking effects in the SSSRT

The structural neural correlates of the different SPE patterns and chunking effects in the sung vs. spoken task in the aphasic and non-aphasic patients were analyzed by correlating them with grey matter volume from voxel-based morphometry data and to the volume and fractional anisotropy of white matter tracts from deterministic tractography data.

Voxel-based morphometry results: In the aphasic group, the larger RE in the sung vs. spoken task correlated with greater grey matter volume in the left posterior temporal [superior temporal gyrus (STG), middle temporal gyrus (MTG)], parietal [postcentral gyrus (postCG), middle occipital gyrus (MOG)], and limbic [parahippocampal gyrus (PHG)] regions as well as in the right frontal [precentral gyrus (preCG)], posterior temporal [inferior temporal gyrus (ITG)], and parietal [inferior parietal lobule (IPL)] regions (see Table 6 and Figure 7A). Similarly, greater grey matter volume in the left posterior temporal (MTG), frontal (preCG), and parietal (postCG, MOG) regions and in the right posterior temporal region (ITG) correlatated with the sung>spoken RE in aphasic compared to non-aphasic patients (see Table 6 and Figure 7B).

Table 6. Significant correlations between grey matter volume and SSSRT performance in Study II (reproduced with permission from the publisher, Leo et al., 2019).

Patients/ Contrast	Condition	Area	MNI coordinates	Cluster size	t-value	Correlation
Aphasia	Sung > spoken RE	Left middle temporal gyrus (BA 21)	-66 -52 1	1008	15.2**	r = .96 p < 0.001
		Left superior temporal gyrus (BA 22)	-69 -41 11			
		Left middle occipital gyrus (BA 19)	-45 -88 13	512	14.5*	r = .95 p < 0.001
		Left postcentral gyrus (BA 3)	-35 -21 44	849	9.4**	r = .95 p < 0.001
		Left parahippocampal gyrus (BA 34)	-13 -19 -22	643	8.2*	r = .95 p < 0.001
		Right inferior temporal gyrus (BA 20, 37)	47 -79 -24	2596	8.5**	r = .96 p < 0.001
		Right cerebellum	53 -65 -25			
		Right inferior parietal lobule (BA 40)	59 -31 34	1184	8.1**	r = .93 p < 0.001
		Right postcentral gyrus (BA 2, 3)	57-22 45			
		Right precentral gyrus (BA 4)	57 -14 36	480	7.3*	r = .95 p < 0.001
Aphasic > Non-aphasic	Sung > spoken RE	Left middle temporal gyrus (BA 37)	-46 -82 15	880	5.7*	n.s.
		Left middle occipital gyrus (BA 19)	-50 -75 0			
		Left middle temporal gyrus (BA 21, 37, 39)	-61 -56 5	782	5.6*	n.s.
		Left precentral gyrus (BA 4)	-41 -23 57	1026	5.0*	r = .74 p = 0.003
		Left postcentral gyrus (BA 3)	-39 -26 57			
		Right inferior temporal gyrus (BA 20, 37)	59 -56 -24	1874	6.0**	r = .55 p = 0.041
		Right cerebellum	47 -70 -23			

All results are thresholded at a whole-brain uncorrected p < 0.001 threshold at the voxel level with a minimal cluster size set to 100 voxels. *p < 0.05 FWE-corrected at the cluster level, **p < 0.005 FWE-corrected at the cluster level. Correlations are partial correlations with 2-tailed p-value controlling for age, sex and total intracranial volume. BA = Brodmann area, RE = recency effect

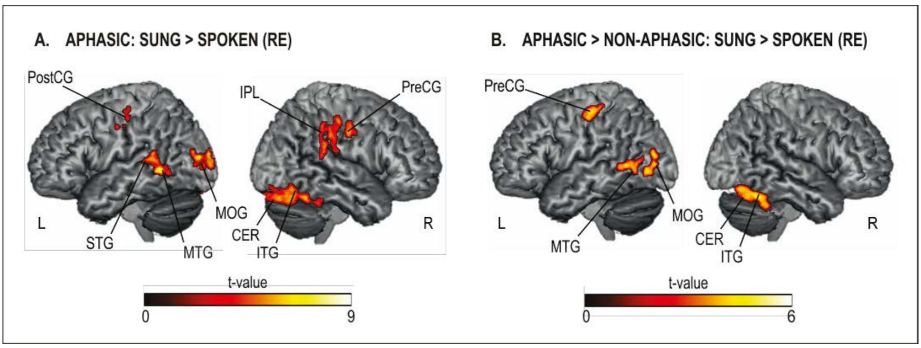

Figure 7 Voxel-based morphometry results showing significant correlations between regional grey matter volume and the sung > spoken recency effect (RE) in the left and right hemispheres (A) within the aphasic group and (B) in a contrast between the aphasic and non-aphasic groups. CER = cerebellum, IPL = inferior parietal lobule, ITG = inferior temporal gyrus, L = left, MOG = middle occipital gyrus, MTG = middle temporal gyrus; PostCG = postcentral gyrus, PreCG = precentral gyrus, R = right, STG = superior temporal gyrus (reproduced with permission from the publisher, Leo et al., 2019).

Deterministic tractography results: In the aphasic patients (Figure 8A), there was a significant correlation between larger RE in the sung than spoken task and larger volume of the right inferior fronto-occipital fasciculus (IFOF; r=0.67, p=.009). In contrast, the non-aphasic patients (Figure 8B) showed a strong correlation between smaller PE in the sung than spoken task and larger volume (r=−0.69, p=.003) and fractional anisotropy (r=−0.63, p=.009) of the left arcuate fasciculus (AF, long segment).

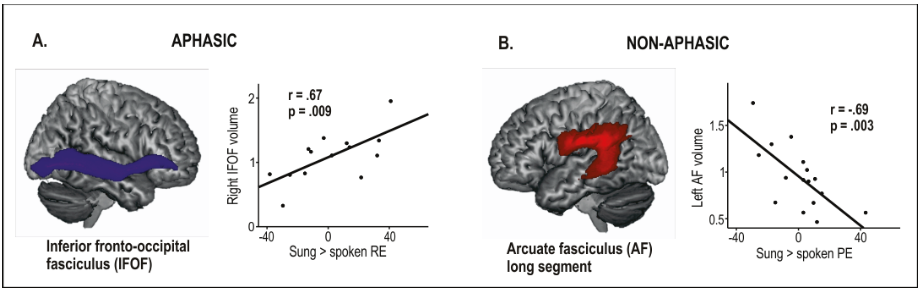

Figure 8 Deterministic tractography results in Study II showing significant correlations between (A) the volume of the right inferior fronto-occipital fasciculus (IFOF) and the sung > spoken recency effect (RE) in aphasic patients and (B) the volume of the left arcuate fasciculus (AF, long segment) and the sung > spoken primacy effect (PE) in non-aphasic patients (reproduced with permission from the publisher, Leo et al., 2019).

4.4 Effects of daily vocal music listening on stroke recovery (Study III)

In order to verify and extend previous results on effects of daily music listening (Särkämö et al., 2008; Baylan et al., 2018, 2019) and to determine the contribution of the singing component of music, the analyses in Study III were performed with both two-group comparisons [VMG and IMG combined to a music group (MG) vs. ABG] and three-group comparisons [VMG vs. IMG vs. ABG] at the acute stage (baseline), from acute to 3-month stage (N = 83), and acute to 6-month stage (N =81).

4.4.1 Behavioral results

At baseline, there were no significant differences between the MG and the ABG or between the VMG, IMG, and ABG in the behavioral outcome measures (Table 7). The longitudinal results of the behavioral outcome measures are shown in Table 7 and Figure 9.

During the follow-up time, the amount of motor, speech, or cognitive rehabilitation received by the patients at 3-month stage and 6-month stage was comparable between the groups (Table 3). However, the frequency of music and audiobook listening differed highly significantly between the groups, both during the intervention and follow-up periods. Listening frequency was higher for music in VMG and IMG than in ABG, and for audiobooks in ABG than in VMG and IMG at 3-month stage ($p<0.001$ in all) and, to a lesser extent, at 6-month stage ($p<0.092$ in all). There were no significant differences in music or audiobook listening between VMG and IMG. The average amount of daily listening to the allocated material was 1.8 hours (SD=0.9), totaling around 100 hours (M=107.9, SD=55.1) over the 2-month intervention period. Even though at group level the listening frequencies followed the study protocol, there was a significant difference between the groups in cross-listening (using own devices to listen to material not part of the protocol: music in ABG, audiobooks in VMG and IMG), indicative of treatment contamination, during the intervention period ($H=42.64$, $p<0.001$), with ABG showing more cross-listening than VMG and IMG ($p<0.001$ in both).

The longitudinal results from the neuropsychological tests (see Table 7 and Figure 9) showed that in the two-group comparisons there were significant Time x Group interactions in verbal memory from acute to 3-month stage [$F(1,71) = 8.70$, $p = 0.004$, $\eta p^2 = 0.109$] and from acute to 6-month stage [$F(1,70) = 14.62$, $p < 0.001$, $\eta p^2 = 0.173$]. This indicates better short-term recovery of language and long-term recovery of verbal memory in the MG than in the ABG. There was also a significant Time x Group interaction in language from acute to 3-month stage [$F(1,61) = 10.40$, $p=0.002$, $\eta p^2 = 0.146$]. Notably, in language there was also a significant Time x Group x Aphasia interaction from acute to 3-month stage [$F(1,61) = 4.53$, $p = 0.037$, $\eta p^2 = 0.069$]. A post hoc subgroup analysis with separate mixed-model ANOVAs within the aphasic and non-aphasic patients showed that language recovered better in the MG than in the ABG only in aphasics [Time x Group: $F(1,20) = 27.63$, $p < 0.001$, $\eta p^2 = 0.580$]. In the mood domain (see Table 8), the MG had lower scores on the POMS Depression [Group: $F(1,70) = 4.02$, $p = 0.049$, $\eta p^2 = 0.054$] and Confusion [Group: $F(1,69)=5.89$, $p=0.028$, $\eta p2=0.079$] scales than the ABG at the 3-month stage. There were no significant effects in the other six POMS scales or in focused attention.

Table 7. Cognitive performance of the patients at different time points in Study III (Sihvonen & Leo et al., *in revision*)

	Stage	MG N=50/49	VMG N=27/26	IMG N= 23	ABG N=33/32	MG/ABG p-value[a]	VMG/IMG/ABG p-value[a]
Verbal memory (score range 0-124)	T0	47.1 (19.7)	43.3 (22.6)	52.3 (14.8)	51.3 (21.8)	0.404	0.305
	T1	62.8 (20.3)	65.3 (20.7)	62.0 (17.9)	64.3 (21.2)	**0.004**	**0.002**
	T2	64.2 (20.0)	63.4 (21.3)	66.8 (17.4)	63.2 (18.9)	**<0.001**	**0.002**
Language skills (score range 0-80)	T0	52.6 (15.1)	48.7 (19.2)	56.8 (6.6)	54.2 (13.1)	0.475	0.431
	T1	62.8 (8.9)	60.1 (10.0)	65.8 (6.3)	59.4 (12.2)	**0.002**	**0.009**
	T2	60.9 (7.0)	58.5 (8.0)	63.8 (4.3)	61.3 (12.0)	0.126	0.297
Focused attention: correct responses (score range 0-90)	T0	75.8 (19.9)	75.6 (20.4)	76.0 (20.1)	76.9 (17.9)	0.748	0.953
	T1	86.3 (8.3)	85.3 (11.2)	87.4 (3.6)	84.1 (11.5)	0.213	0.303
	T2	88.4 (2.1)	88.4 (1.8)	88.5 (2.4)	85.0 (7.8)	0.139	0.303
Focused attention: reaction times (s)	T0	4.5 (4.5)	3.8 (1.9)	5.1 (6.0)	4.6 (2.5)	0.451	0.189
	T1	3.4 (3.7)	2.7 (1.0)	4.2 (5.2)	3.0 (1.4)	0.523	0.863
	T2	3.6 (4.4)	2.7 (1.4)	4.5 (6.1)	2.9 (1.3)	0.376	0.717

Data are mean (SD). Higher scores indicate better outcome. Significant group differences are shown in bold. Abbreviations: ABG = Audio book group, IMG = Instrumental music group, MG = Music group (VMG & IMG combined), T0 = baseline (acute), T1 = 3-month stage, T2 = 6-month stage, VMG = Vocal music group. The p values are from Group main effect in univariate ANOVA at T0 and from Time x Group interaction in mixed-model ANOVA at T1 (T0 vs T1) and T2 (T0 vs T2).

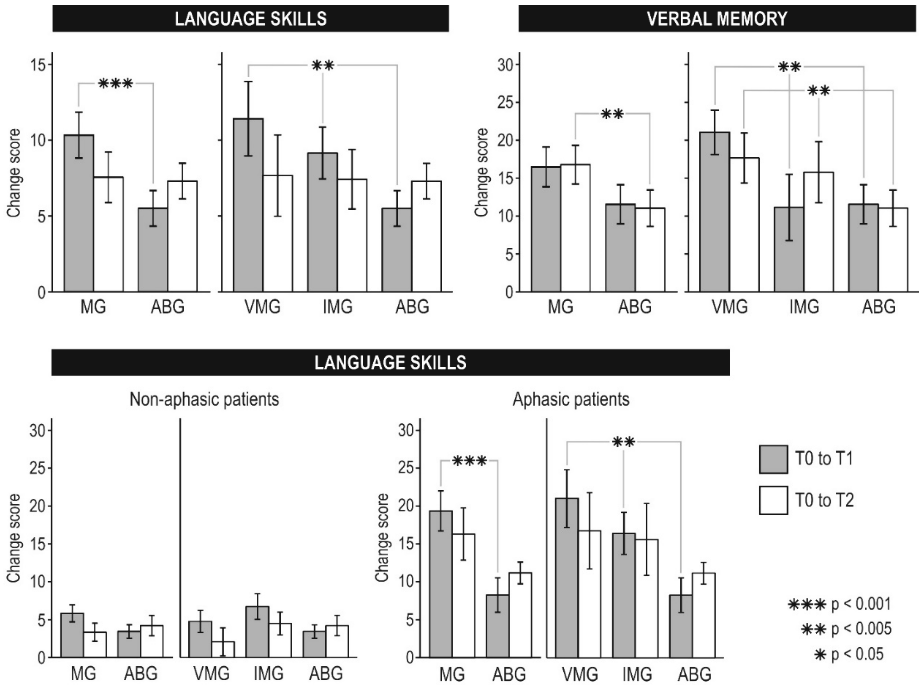

Figure 9 Changes in language skills and verbal memory and (mean ± SEM) from acute to 3-month (grey bars) and 6-month (white bars) post-stroke stage in the three groups. Upper panel shows results across all patients and lower panel within patients without aphasia and with aphasia. The significance of the Time x Group interaction is shown with asterisks. ABG = Audio book group, IMG = Instrumental music group, MG = Music group (VMG & IMG combined), T0 = baseline (acute), T1 = 3-month stage, T2 = 6-month stage, VMG = Vocal music group (Sihvonen & Leo et al., *in revision*)

In the three-group comparisons, there were significant Time x Group interactions in verbal memory from acute to 3-month stage [$F(2,67) = 6.96$, $p = 0.002$, $\eta p^2 = 0.172$] and from acute to 6-month stage [$F(2,66) = 7.10$, $p = 0.002$, $\eta p^2 = 0.177$]. Bonferroni-corrected post hoc testing showed that verbal memory improved more in the VMG than in the IMG from acute to 3-month stage ($p = 0.015$) and more in the VMG than in the ABG from acute to 6-month stage ($p = 0.046$). In language, there were significant Time x Group [$F(2,58) = 5.12$, $p = 0.009$, $\eta p^2 = 0.150$] and Time x Group x Aphasia [$F(2,58) = 3.42$, $p = 0.039$, $\eta p^2 = 0.105$] interactions from acute to 3-month stage. A post hoc subgroup analysis with separate mixed-model ANOVAs within aphasics and non-aphasics showed that the Time x Group interaction was significant only in the aphasics [$F(2,18) = 12.82$, $p<0.001$, $\eta p^2 = 0.588$]. Bonferroni-corrected post hoc testing within the aphasics showed that language recovery was better in the VMG than in the ABG ($p = 0.002$). There were no significant effects in the POMS scales or in focused attention.

In summary, these results suggest that compared to audiobooks, daily music listening enhances the recovery of language (especially in aphasic patients) and verbal memory and reduces negative mood and that these rehabilitative effects are strongest for vocal music.

Table 8. Mood of the patients at different time points in study III (Sihvonen & Leo et al., *in revision*)

	Stage	MG N=50/49	VMG N=27/26	IMG N= 23	ABG N=33/32	MG / ABG p-value	MG / IMG / ABG p-value
POMS Tension (score range 0-16)	T0	4.2 (3.3)	3.8 (3.2)	4.4 (3.2)	4.2 (3.6)	0.933	0.684
	T1	3.5 (2.7)	3.9 (3.3)	2.9 (2.2)	3.0 (2.8)	0.227	0.597
	T2	3.1 (2.8)	3.5 (3.4)	2.6 (2.0)	3.3 (2.8)	0.805	0.875
POMS Depression (score range 0-28)	T0	6.2 (5.9)	5.3 (5.7)	7.0 (6.3)	6.7 (6.9)	0.407	0.560
	T1	3.5 (3.7)	3.8 (3.9)	3.0 (2.6)	5.7 (5.4)	**0.049**	0.103
	T2	3.8 (4.6)	4.9 (5.6)	3.5 (3.0)	5.1 (5.9)	0.484	0.612
POMS Irritability (score range 0-28)	T0	5.1 (5.2)	4.1 (5.6)	6.1 (4.6)	4.4 (6.0)	0.992	0.753
	T1	4.7 (4.1)	5.8 (4.4)	3.4 (3.5)	5.4 (6.9)	0.182	0.164
	T2	4.5 (4.7)	5.9 (5.5)	2.9 (3.1)	5.2 (5.2)	0.911	0.401
POMS Vigor (score range 0-24)	T0	9.5 (5.7)	10.9 (5.5)	8.1 (5.1)	8.8 (5.2)	0.256	0.152
	T1	12.8 (6.6)	13.7 (6.9)	12.0 (6.4)	11.3 (5.9)	0.521	0.387
	T2	13.1 (6.0)	13.0 (6.6)	13.3 (5.6)	11.6 (6.3)	0.746	0.522
POMS Fatigue (score range 0-12)	T0	5.6 (2.7)	5.5 (2.7)	5.5 (2.8)	5.0 (2.9)	0.677	0.822
	T1	4.1 (2.5)	4.1 (2.3)	4.0 (2.7)	4.7 (3.1)	0.951	0.913
	T2	3.7 (2.9)	4.1 (3.0)	3.3 (2.9)	3.5 (3.1)	0.403	0.567
POMS Inertia (score range 0-12)	T0	2.9 (2.3)	2.6 (2.0)	3.0 (2.5)	3.0 (2.9)	0.649	0.813
	T1	2.5 (2.5)	2.6 (2.7)	2.3 (2.4)	3.2 (2.4)	0.309	0.639
	T2	2.7 (2.7)	2.8 (2.7)	2.6 (2.9)	2.9 (3.1)	0.679	0.405
POMS Confusion (score range 0-20)	T0	7.3 (4.3)	7.4 (4.3)	7.2 (4.6)	7.4 (4.5)	0.664	0.884
	T1	3.1 (3.3)	3.5 (3.7)	2.4 (2.8)	5.1 (3.4)	**0.028**	0.096
	T2	3.3 (3.3)	3.8 (3.7)	2.8 (2.8)	3.5 (3.2)	0.824	0.992
POMS Forgetfulness (score range 0-12)	T0	4.3 (2.7)	4.2 (2.9)	4.1 (2.6)	4.2 (2.5)	0.903	0.796
	T1	3.2 (2.6)	3.4 (2.6)	3.0 (2.6)	3.5 (2.4)	0.437	0.768
	T2	3.3 (2.9)	3.4 (3.0)	3.0 (3.0)	3.2 (2.7)	0.999	0.813

Data are mean (SD). Lower scores indicate better outcome, except in the Vigor scale. Significant group differences are shown in bold. Abbreviations: ABG = Audio book group, IMG = Instrumental music group, MG = Music group (VMG & IMG combined), POMS = Profile of Mood States, T0 = baseline (acute), T1 = 3-month stage, T2 = 6-month stage, VMG = Vocal music group. The p values are from Group main effect in univariate ANOVA at T0, T1, and T2.

4.4.2 Voxel-based morphometry results

The longitudinal voxel-based morphometry results are presented in Figure 10 and Table 9. In the two-group comparisons, there was a larger grey matter volume increase in the MG than in the ABG in two clusters in left anterior temporal [superior (STG), middle (MTG), and inferior (ITG) temporal gyrus] and left parietal [precuneus, angular gyrus (AG), inferior parietal lobule (IPL), middle occipital gyrus (MOG)] areas across all patients from acute to 6-month stage (Figure 10A). In contrast, within the aphasic patients, grey matter volume increased more in the MG than in the ABG in one cluster in right posterior temporal (MTG) and parietal (precuneus, AG, MOG) areas (Figure 10B), and the increased grey matter volume in this cluster correlated with the improvement of language from acute to 3-month ($r = .57$, $p = 0.042$) and 6-month ($r = .71$, $p = 0.008$) stage.

In the three-group comparisons, grey matter volume increased more in the VMG than in the ABG in one cluster in left anterior temporal (STG, MTG, ITG) areas across all patients from acute to 6-month stage (Figure 10A). In aphasics, white matter volume increased more in the VMG than in the ABG in one cluster comprising right medial parieto-occipital [lingual gyrus (LG), cuneus, MOG] areas (Figure 10C), and the increased white matter volume in this cluster correlated with the improvement of verbal memory and language from acute to 3-month ($r = .72$, $p = 0.004$ and $r = .80$, $p < 0.001$) and 6-month ($r = .68$, $p = 0.005$ and $r = .56$, $p = 0.024$) stage. In aphasics, there was also larger grey matter volume increase in the IMG than in the ABG in one cluster in right temporoparietal (MTG, MOG) areas (Figure 3B), and the increased grey matter volume in this cluster correlated with the improvement of language from acute to 6-month stage ($r = .64$, $p = 0.040$).

In summary, the voxel-based morphometry results suggest that the positive effects of music listening on language and verbal memory recovery are coupled with increased structural neuroplasticity, indicated by changes in grey matter volume and white matter volume, especially in left temporal and parietal areas in all patients and in right medial parietal areas in aphasic patients.

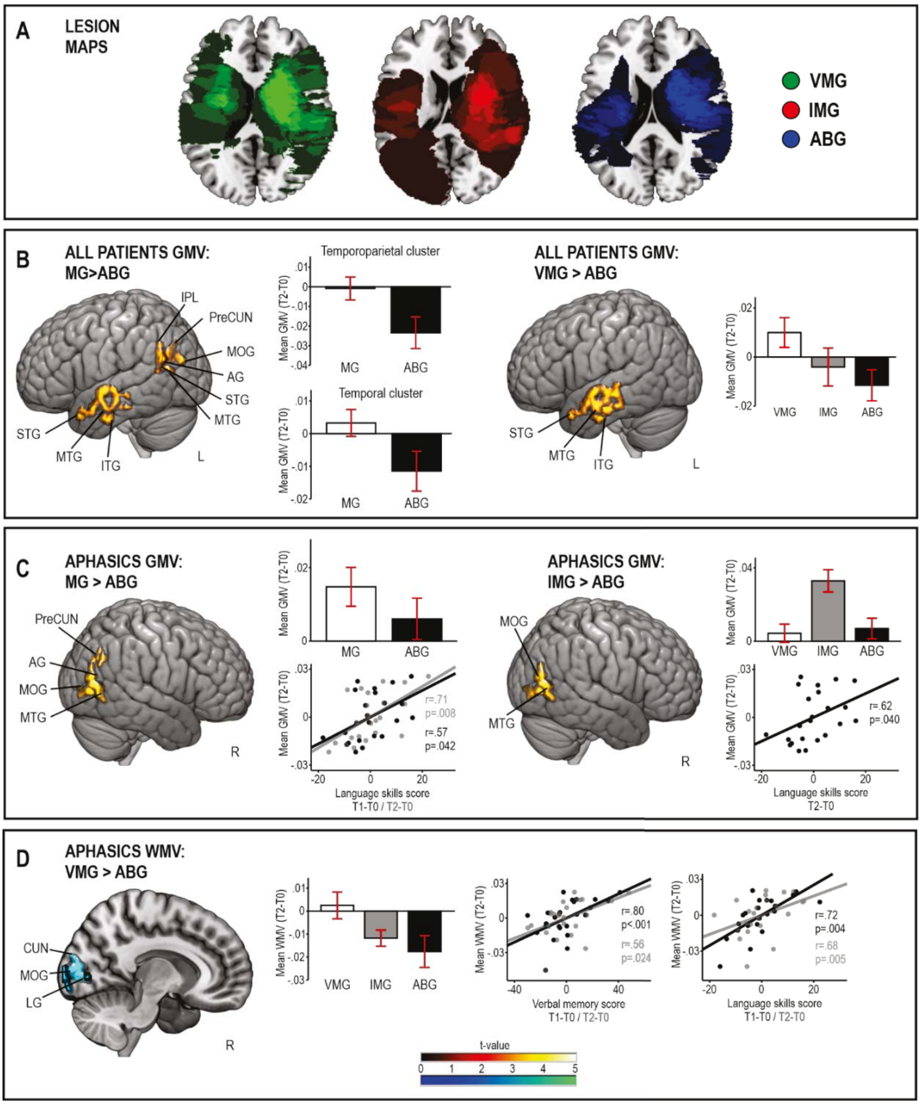

Figure 10 Pooled voxel-based morphometry results from the Helsinki and Turku studies (N = 75). Panel A: Lesion overlay maps of the three groups. Panels: B-D: Significant group differences in voxel-based morphometry VBM from acute to 6-month stage in (B) grey matter volume (GMV) across all patients in two-group (left) and three-group (right) comparisons, (C) GMV within aphasic patients in two-group (left) and three-group (right) comparisons, and (D) white matter volume (WMV) within aphasic patients three-group (left) comparison. Results are shown at p<0.005 (uncorrected) with ≥100 voxels of spatial extent. Only clusters surviving a FWE-corrected p<0.05 threshold are shown and labelled. Group differences are shown with histograms (mean ± SEM). Correlations to change in language / verbal memory are shown with scatter plots. ABG = Audio book group, IMG = Instrumental music group, MG = Music group (VMG & IMG combined), T0 = baseline (acute), T1 = 3-month stage, T2 = 6-month stage, VMG = Vocal music group. Anatomical abbreviations: AG = angular gyrus, CUN = cuneus, IPL = inferior parietal lobule, ITG = inferior temporal gyrus, LG = lingual gyrus, MOG = middle occipital gyrus, MTG = middle temporal gyrus, PreCUN = precuneus, STG = superior temporal gyrus (Sihvonen & Leo et al., *in revision*).

Table 9. Grey and white matter volume changes (acute to 6-month stage) in voxel-based morphometry analyses in Study III (Sihvonen & Leo et al., *in revision*).

Patients	Group contrast	Volume type	Area	MNI coordinates	Cluster size	FWEc	t-value	Correlation between volume and cognitive change: T1-T0 LAN	T1-T0 MEM	T2-T0 LAN	T2-T0 MEM
All	MG>ABG	GMV	Left middle temporal gyrus (BA 21)	-57 -8 -5	5119	3333	4.48**	n.s.	n.s.	n.s.	n.s.
			Left superior temporal gyrus (BA 22, 38)	-54 -6 -5							
			Left inferior temporal gyrus (BA 20)	-54 -27 -20							
			Left superior temporal gyrus (BA 39)	-59 -61 27	3333		3.86*	n.s.	n.s.	n.s.	n.s.
			Left middle temporal gyrus (BA 39)	-53 -70 25							
			Left angular gyrus (BA 39)	-49 -68 31							
			Left precuneus (BA 39)	-39 -70 34							
			Left inferior parietal lobule (BA 40)	-54 -57 39							
			Left middle occipital gyrus (BA 19)	-39 -78 21							
	VMG>ABG	GMV	Left superior temporal gyrus (BA 22, 38)	-57 -7 -5	7070	7070	4.33**	n.s.	n.s.	n.s.	n.s.
			Left middle temporal gyrus (BA 21)	-62 -16 -1							
			Left inferior temporal gyrus (BA 20)	-62 -10 -18							
Aphasics	MG>ABG	GMV	Right middle temporal gyrus (BA 37, 39)	44 -74 21	3120	3120	5.87**	n.s.	n.s.	r=.57 p=.042	r=.71 p=.008
			Right middle occipital gyrus (BA 19)	49 -74 12							
			Right angular gyrus (BA 39)	50 -66 29							
			Right precuneus (BA 19)	32 -66 42							
	VMG>ABG	WMV	Right lingual gyrus	10 -79 0	2498	2498	6.17*	r=.80 p<.001	r=.56 p=.024	r=.72 p=.004	r=.68 p=.005
			Right cuneus	8 -87 9							
			Right middle occipital gyrus	9 -96 12							
	IMG>ABG	GMV	Right middle temporal gyrus (BA 37, 39)	44 -74 21	2893	2893	6.16*	n.s.	n.s.	n.s.	r=.62 p=.040
			Right middle occipital gyrus (BA 19)	47 -82 4							

All results are thresholded at a whole-brain uncorrected $p<0.005$ threshold at voxel level with a minimal cluster size of 100 voxels.
*$p<0.05$ FWE-corrected at the cluster level
**$p<0.005$ FWE-corrected at the cluster level
FWEc is the minimum number of voxels for a cluster to be significant at the FWE-corrected $p<0.05$ level, according to SPM standard cluster-level correction based on random field theory and cluster-forming threshold of $p<0.005$.
ABG = Audiobook group, BA = Brodmann area, GMV = grey matter volume, IMG = Instrumental music group, LAN = language skills, MEM = verbal memory, MG = Music group, VMG = Vocal music group, WMV = white matter volume

4.4.3 fMRI functional connectivity results

The music-induced grey matter volume and white matter volume changes were located primarily within the posterior and temporal parts of the default mode network (Laird et al., 2009; Raichle, 2015), a large-scale brain system that participates in internal cognition (Buckner, Andrews-Hanna & Schacter, 2008), which has recently been linked also to episodic or verbal memory (Staffaroni et al., 2018; Zhang, Andreano, Dickerson, Touroutoglou, 2019). In order to explore if changes in the default mode network could underlie the cognitive benefits and structural neuroplasticity induced by the vocal music listening, functional connectivity analyses were performed on longitudinal (acute and 6-month) fMRI data from the Turku patients acquired at rest, with no auditory stimuli (resting-state fMRI) and during a task (task-fMRI) involving listening to 15-second excerpts of well-known Finnish songs with sung lyrics (Vocal), without sung lyrics (Instrumental), and well-known Finnish poems (Speech).

In the resting-state fMRI independent component analysis, which focused on the spatial component of each brain network, the VMG showed functional connectivity between the whole default network and left temporal (STG, MTG) areas more than ABG or IMG. Furthermore, the VMG showed increased functional connectivity more than IMG between the right temporal (STG, Heschl's gyrus) areas and rest of the default mode network from acute to 6-month stage (Figure 11 and Table 10). In the task-fMRI, there was a significant Time x Group interaction from acute to 6-month stage [$F(1,25) = 3.73$, $p = 0.038$] in whole network-level default mode network engagement in the Vocal condition (Figure 11). Bonferroni-corrected post hoc tests showed a larger increase in the VMG than in the ABG ($p = 0.041$). No significant effects were observed in the Instrumental and Speech conditions. Correlation analyses showed that in the VMG patients, the increased resting-state functional connectivity between the different clusters of the default mode network and the left STG/MTG correlated with the improvement of language from acute to 3-month stage ($r = .78$, $p = 0.040$) and verbal memory from acute to 6-month stage ($r = .66$, $p = 0.040$). In addition, in resting state fMRI connectivity between the left temporal regions and the default mode network also correlated with the mean default mode network engagement in during the Vocal condition: the greater the default mode network engagement after 6 months while listening to vocal music, the more functionally connected the left temporal lobe is with the default mode network ($r = .60$, $p = 0.038$).

In summary, the fMRI results suggest that the vocal music listening induces also functional neuroplasticity changes, indicated by resting-state and task-related functional connectivity increase in the default mode network. These are linked to the long-term positive effects of vocal music on language and memory recovery and overlap partially with the left temporal and medial parietal regions where structural neuroplasticity changes were observed in the voxel-based morphometry analysis.

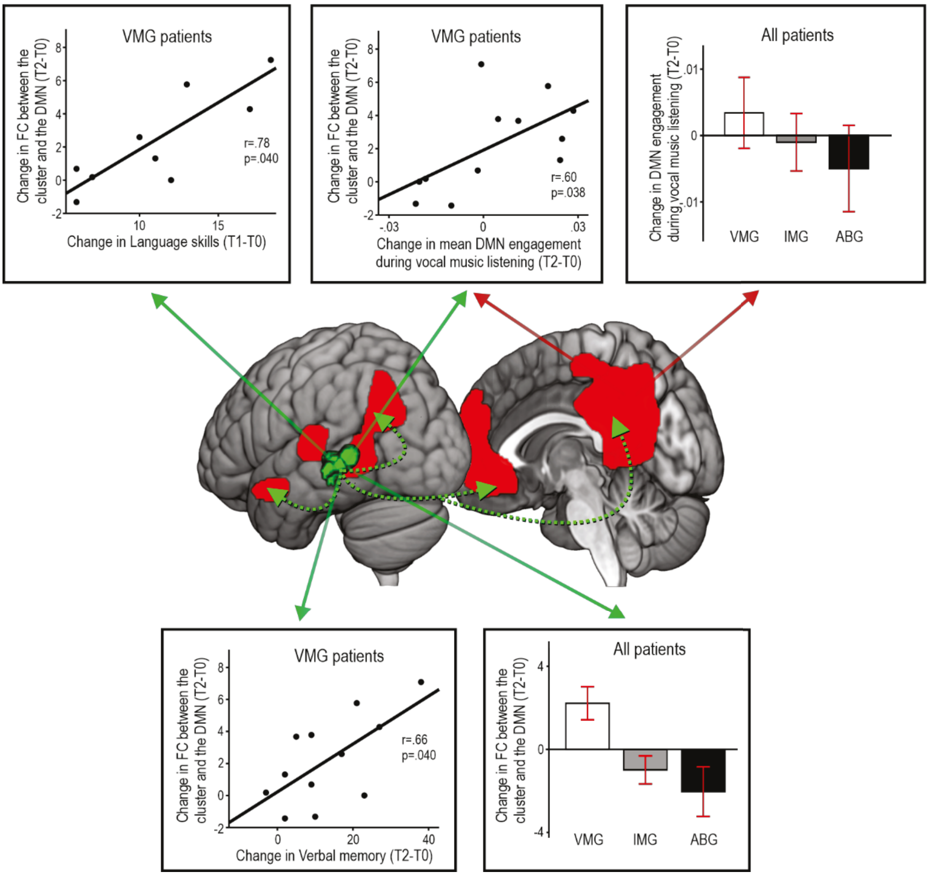

Figure 11 fMRI results on functional connectivity from the Turku study (N=35). Significant group differences in functional connectivity (FC) from acute (T0) to 6-month (T2) stage between the default mode network (DMN) and the left temporal (STG/MTG) areas in the resting-state condition (cluster shown in green-black color scale) and the mean engagement of the DMN during the Vocal music listening condition (DMN illustrated in red color). Spatial results in the resting-state condition are at $p<0.005$ (uncorrected) with ≥100 voxels of spatial extent, and only the cluster surviving a FWE-corrected $p<0.05$ threshold is shown. Group differences are shown in histograms (mean ± SEM), and correlations within the VMG (N = 12) are shown in scatter plots. The histograms / scatter plots showing results in the resting-state condition are marked with green solid arrows and those showing results in the vocal music listening condition are marked with red solid arrows. Dashed green arrows illustrate FC between the left temporal cluster and the other parts of the DMN. ABG = Audio book group, IMG = Instrumental music group, T0 = baseline (acute), T1 = 3-month stage, T2 = 6-month stage, VMG = Vocal music group. (Sihvonen & Leo et al., *in revision*).

Table 10. Functional connectivity changes (acute to 6-month stage) in resting-state fMRI analyses in Study III (Sihvonen & Leo et al., *in revision*).

Group contrast	Area	MNI coordinates	Cluster size	FWEc	t-value
VMG>ABG	Left superior temporal gyrus Left middle temporal gyrus	-44 -34 4 -56 -40 4	342	342	4.52*
VMG>IMG	Left middle temporal gyrus Left superior Temporal gyrus	-46 -48 3 -45 -31 2	338	296	4.98*
	Right Heschl's gyrus Right superior temporal gyrus	46 -22 5 48 -18 4	296		4.48*

All results are thresholded at a whole-brain uncorrected p<0.005 threshold at voxel level with a minimal cluster size of 100 voxels.
*p<0.05 FWE-corrected at the cluster level
FWEc is the minimum number of voxels for a cluster to be significant at the FWE-corrected p<0.05 level, according to SPM standard cluster-level correction based on random field theory and cluster-forming threshold of p<0.005.
ABG = Audiobook group, IMG = Instrumental music group, LAN = language skills, MEM = verbal memory, VMG = Vocal music group

5 DISCUSSION

The three studies reported in this thesis explored the mnemonic and rehabilitative effects of songs after stroke by determining if sung presentation modality can aid in the learning and recall of novel verbal material (Study I) and which cognitive and neural mechanisms underlie this effect (Study II) and if daily music listening, especially of vocal music, can enhance verbal, cognitive, emotional, and neural recovery (Study III).

The main findings of the thesis were:

Study I: Stroke patients benefit from sung melody as a mnemonic aid in the learning and recall of novel verbal material (narrative stories) 6 months post-stroke. Sung melody aids verbal learning especially in patients with mild aphasia.

Study II: The mnemonic effect of songs is underpinned by a number of cognitive and neural mechanisms, which differ depending on level of aphasia. Sung presentation supports, in non-aphasic patients, more stable recall of across the story (smaller SPE), mediated by the left dorsal pathway (AF). In aphasic patients sung presentation enhances chunking and better recall of the last parts of the story (larger RE), mediated by bilateral frontal, temporal, and parietal areas and the right ventral pathway (IFOF).

Study III: Daily music listening enhances the recovery of language (especially in aphasic patients) and verbal memory as well as reduces negative mood. Vocal music has the strongest rehabilitative effect on both language and verbal memory. These effects are coupled with structural neuroplasticity changes in left temporal and parietal areas as well as in right medial parietal areas specifically in aphasic patients and with functional connectivity changes in the default mode network.

5.1 Songs as a verbal learning and memory aid after stroke

The results of Study I show that patients benefited from the sung melody in recalling novel stories at the 6-month post-stroke stage. The same effect was not seen at the acute stage; in fact, aphasic patients performed slightly worse in the sung than spoken task at the acute stage (see Figure 3). This lack of positive effect at the acute stage may be associated with higher severity of verbal and cognitive deficits and fatigue. Post-stroke fatigue is highly prevalent at the acute stage, affecting 25-75% of patients (Pihlaja, Uimonen, Mustanoja, Tatlisumak & Poutiainen, 2014). Deficits in general attentional shifting and slowness of information processing, which are also common early after stroke (Ramsey et al., 2017), may also have made the parallel encoding of the linguistic and musical information difficult to the patients. In this case, the musical element may bring an additional cognitive demand to the task, rather than providing an aid to memory.

At the 6-month stage, patients with at least mild aphasia showed better learning of the sung stories than the spoken stories, starting from the second learning trial onwards, indicating that the effect builds up with repetition and rehearsal. In aphasic patients, the sung > spoken effect increased considerably from acute to 6-month stage and was evident in all three learning trials and in the delayed recall trial in those aphasics who did not have concurrent memory impairment. In addition, at the 6-month stage, also amusic patients and memory-impaired patients showed the sung > spoken effect in delayed recall and memory-impaired patients also in the 2nd learning trial.

Thus far, evidence from previous studies comparing the learning of novel verbal material in sung vs. spoken modality has been to some extent controversial. Overall, studies in healthy subjects suggest that novel sung verbal material can be learned and recalled better than spoken material, but typically this effect requires that the task involves connected verbal material (phrases, sentences, stories) and multiple learning trials and that the melody is simple and repetitive across verses, and accentuates the surface characteristics of the text. In addition, the tempo of the singing should be natural, which is about 30-50% slower than in speech, and the spoken version should have normal prosodic structure (Wallace, 1994; McElhinney & Annett, 1996; Kilgour et al., 2000; Purnell-Webb & Speelman, 2008; Racette & Perez, 2007). The SSSRT fits closely within this framework, as (i) the novel stories were arranged in five verses and were repeated over several learning trials, (ii) their content was concrete and narrative, with a clear storyline, (iii) the melodies were structurally simple and repeated across verses, and (iv) the singing was in natural tempo (slower than speech).

These four stimulus and task-related factors make the integration of the text and the melody possible, reducing the likelihood of a dual-task situation where the melody and text are treated separately and are thereby more challenging for attention (Wallace, 1994; Kilgour et al., 2000; Racette & Peretz, 2007; Ferreri & Verga, 2016). At the neural level, the effect is most likely related to the bilaterality and the breadth of the temporal, frontal, and limbic networks, which are engaged in the processing of songs (Callan et al., 2006; Brattico et al., 2011; Alluri et al., 2013) and, consequently, to the preservation of this wide-spread network after focal brain damage. For example, using fMRI, Sihvonen et al. (2017b) showed that in stroke patients listening to sung music activated more extensive temporal and frontal regions than listening to instrumental music (Sihvonen et al., 2017b). Singing-based training has been observed to be beneficial both for enhancing speech production in aphasia (Schlaug et al., 2008; van der Meulen et al., 2014; Zumbansen, Peretz, & Hébert, 2014b) and also singing production and music perception in amusia (Anderson, Himondes, Wise, Welch & Stewart, 2012; Wilbiks, Vuvan, Girard, Peretz & Russo, 2016).

Together, these factors most likely contribute to the better recall of sung verbal material in neurological patients. The sung > spoken effects observed in Study I are in line with similar findings from studies of persons with multiple sclerosis (Thaut et al. 2014) and Alzheimer's disease (Simmons-Stern et al., 2010; Moussard et al., 2014; Palisson et al., 2015). However, previous studies with subjects with aphasia have not found differences in the recall of novel sung and spoken material (Hébert et al., 2003; Racette et al., 2006). Comparison between these two studies with the Study I reveals some pivotal methodological differences. First, the group of patients with aphasia was larger and the severity of aphasia was milder in Study I whereas previous studies have included 1-8 patients with moderate to severe aphasia (Hébert et al., 2003; Racette et al., 2006). Second, the stories in Study I were more concrete and narrative in their content, whereas the previous studies used unfamiliar song lyrics (Hébert et al., 2003; Racette et al., 2006). Third, the task of Study I was cognitively more demanding and the spoken comparison did not include a melodic cue. These factors could have possibly accentuated the sung > spoken effect at the 6-month stage in Study I.

Study I has some methodological issues, which need to be discussed. First of all, in the SSSRT, the word presentation rate was slower in the sung than spoken task. The duration of the sung versions was approximately 50% longer than the spoken versions. It is possible that the slower rate potentially facilitated the recall of the sung stories, since slower rate gives more time to encode the lyrics. Essentially, it would have been possible to control for this effect by including a slowed down spoken version (or alternatively a speeded up sung version) as a control task. This was not done, however, since manipulation of the speed in either the sung or the spoken version would have made the stimuli sound unnatural and artificial. This itself could have attracted the patients' attention away from the content of the text, thereby introducing an additional unwanted element to the task, as well as made the task less ecologically valid. The slower production rate is an innate feature of singing. It is also an essential component in the rehabilitative use of singing in many neurological communication impairments including aphasia (Wan et al., 2010). Secondly, even though the verbal content of the sung and spoken versions of the SSSRT was counter-balanced, the presentation order of the tasks was not: the spoken task was always performed first. This was done intentionally to avoid a potential carry over effect of the sung task on the spoken task, like humming or imaging the melody in mind during the spoken task, when performing the sung task first. This would have been possible since both of the stories were designed to fit to the melody. The fixed presentation order raises the question whether the better recall of the sung task could be, at least partly, due to a general practice effect. The SSSRT was relatively long and part of a larger neuropsychological testing battery, and the sung task was the final test of the session. Because of this, it is likely that the potential bias arising from the fixed order was offset by the increasing fatigue when performing the tasks. Therefore, it is unlikely that the presentation order had significant effect on observed results.

5.2 Cognitive mechanisms underlying the sung > spoken recall effect

The findings of Study I raised a question whether, in addition to the stimulus and task-related factors discussed above, the sung > spoken effect could additionally be underpinned by cognitive mechanisms influencing how the verbal information is processed. Specifically, Study II sought to determine (i) if there were differences between the sung and spoken tasks in the involving memory processes, indicated by the serial position effect (SPE) and chunking, and (ii) if these effects would differ in patients with and without aphasia. Furthermore, Study II also sought to uncover the neural correlates of the observed effects.

The results of Study II showed that there were indeed divergences in SPE and chunking in the two tasks, and that these differed between the non-aphasic and aphasic patients. Non-aphasic patients showed a classic SPE pattern of primacy (PE) and recency (RE) effects in the spoken task, but no observable SPE in the sung task. Verse-level comparison between the tasks revealed that the non-aphasic patients recalled the middle verse (V3) better in the sung task than the spoken task, which resulted in a more stable recall performance in the sung task. This finding is in line with previous studies in healthy subjects, which have reported a smaller SPE for the learning of digits when presented in the sung than spoken format (Silverman, 2007) and also when recalling familiar song lyrics (Maylor, 2002). At the neural level, the deterministic tractography results showed that the better stability of recall in the sung task (resulting a smaller PE) compared to the spoken task correlated with the volume and integrity (fractional anisotropy) of the left arcuate fasciculus (AF). The left AF is thought to map sensory targets in posterior temporal areas to motor programs that are coded in Broca's area (Hickok & Poeppel, 2007). Forming the dorsal "perception-action pathway", the left AF is an important tract for verbal learning in both children (Leroy et al., 2011; Su et al., 2018) and in adults (López-Barroso et al., 2013; Thiebaut de Schotten, Cohen, Amemiya, Braga, & Dehaene, 2014). It also is considered essential for the memory rehearsal process in verbal working memory (Buchsbaum, Olsen, Koch, & Berman, 2005). The left AF forms a key part of the dual-stream pathway for singing production (Loui, 2015) and it shows neuroplastic changes as a result of singing training (Halwani, Loui, Rüber, & Schlaug, 2011). It is possible that the cues provided by the repetitive melodic structure and the slower presentation rate in the sung task made it easier for covert rehearsal and less demanding for the working memory, resulting in less recruitment of the left dorsal pathway than in the spoken task.

In contrast, the aphasic patients showed a different SPE pattern, with a PE in both tasks but a RE only in the sung task. The RE was also larger in the sung than spoken task in aphasic compared to non-aphasic patients. This was attributable primarily to aphasic patients' better recall of the last verse (V5) in the sung task. Furthermore, compared to non-aphasic patients, the aphasic patients recalled longer chunks in the sung task than spoken task. These results are in line with the hypothesis that the sung melody helps to combine words and links subsequent verses together in memory, enabling chunking (Ferreri et al., 2015; McElhinney & Annett, 1996; Wallace, 1994), and that the repetition of the same melody across the verses builds the mnemonic effect in the sung task. Once the melody is learned, it makes the last part of the story more prominent and easily accessible for recall, thereby enhancing the RE. Previous studies in patients with non-fluent aphasia revealed that they have relatively normal RE and clearly smaller PE in recalling serial verbal material. This appears to be related to the reduced verbal memory span and difficulties in covert rehearsal typical of aphasia (Ivanova et al., 2018; Jefferies, Hoffman, Jones & Ralph, 2008; Ostergaard & Meudell, 1984). Accordingly, in Study II, it is possible that the aphasic patients did not benefit from the sung melody as an aid in covert rehearsal in working memory, which, as discussed above, presumably helped the non-aphasic patients, who also had less severe memory deficits. Instead, in aphasic patients the sung > spoken effect seemed to be driven more by stimulus-specific factors, especially the repetitive melody and rhythm of the song, which enabled chunking and made the last verse most salient in short-term storage. Moreover, recalling the sung melody probably also acted as a contextual cue for the patients when attempting to retrieve the lyrics from memory. This latter idea is supported by one study in healthy subjects, in which a stronger RE in delayed recall of a word list was found when same environmental cues, including background music played during the task, were present in both the initial encoding and the delayed recall situation (Isarida & Isarida, 2006).

In the voxel-based morphometry analyses, the larger RE in the sung task compared to the spoken task in the aphasic patients was associated with larger grey matter volume in the left posterior temporal (STG, MTG), parietal (postCG, MOG), and limbic (PHG) regions and also in right frontal (preCG), posterior temporal (ITG), and parietal (IPL) regions. Excluding the right frontoparietal areas, these same regions emerged also in the contrast between the aphasic and non-aphasic groups. Overall, these regions were different from those associated with the general sung > spoken learning effect in aphasics, which was seen primarily in left prefrontal areas (IFG, MFG, anterior cingulate) and bilateral superior parietal and striatal areas (Särkämö & Sihvonen, 2018). The voxel-based morphometry results of Study II are, however, in line with previous neuroimaging studies in healthy subjects and memory-impaired neurological patients, which have linked the RE specifically to the IPL (Buchsbaum et al., 2011; Innocenti et al., 2013), MTG and ITG (Düzel et al., 1996; Spalletta et al., 2016; Staffaroni et al., 2017), and hippocampal area (Spalletta et al., 2016). Based on previous neuroimaging studies of singing in healthy subjects, the MTG/STG, preCG, and cerebellar regions are probably associated with the perceptual processing of lexical/phonological and melodic features and with vocal-motor processing (Callan et al., 2006; Méndez Orellana et al., 2014; Özdemir et al., 2006; Salmi et al., 2017; Schön et al., 2010; Segado, Hollinger, Thibodeau, Penhune & Zatorre, 2018). The right IPL, in turn, has been identified as a crucial hub for higher-level analysis of melodic (Royal et al., 2016) and rhythmic (Konoike et al., 2012) structure of music.

In neuroimaging studies, aphasia recovery has been attributed to the functioning of both left and right hemisphere language networks (Forkel et al., 2014; Saur et al., 2006). For example, melodic intonation therapy has been found to increase functional activation during speech production in frontotemporal auditory-motor and language areas in either the left or right hemisphere or bilaterally (Belin et al., 1996; Breier et al., 2010; Laine et al., 1993b; Schlaug et al., 2008). In diffusion tensor imaging studies (Wan et al., 2014; Zipse, Norton, Marchina, & Schlaug, 2012), structural neuroplasticity changes evoked by melodic intonation therapy have mainly been reported in right frontotemporal tracts, including the AF and the uncinate fasciculus (UF). In Study II, the larger RE in the sung than spoken task correlated with larger volume of the right IFOF in aphasic patients. Part of the ventral stream, the IFOF has a role in language processing (Dick & Tremblay, 2012) and cognition (Cremers et al., 2016), especially in working memory and learning (Chiou et al., 2016; Krogsrud et al., 2018) and in music perception (Sihvonen et al., 2017c). The posterior termination branches of the IFOF extend to the posterior inferior temporal and parietal areas (Duffau, Herbet, & Moritz-Gasser, 2013; Sarubbo, De Benedictis, Maldonado, Basso, & Duffau, 2013), close to the ITG and IPL clusters that were linked to the RE in our aphasic patients and in previous imaging studies (see above). Given its role in multimodal integration and semantic processing (Sarubbo et al., 2013), it is plausible that in our sung task the IFOF carries the melodic and rhythm information from posterior temporal and parietal areas to prefrontal areas (IFG, MFG) where this perceptual information is integrated with the verbal content of the story and processed in working memory.

5.3 Vocal music listening as a rehabilitation tool for language and verbal memory after stroke

Previous findings about the positive effects of daily music listening on cognitive and neural recovery after stroke (Särkämö et al., 2008, 2010, 2014; Baylan et al., 2018, 2019) raise the question which specific factors in music underlie its rehabilitative effect. Study III focused on determining the role of the sung (lyrical) component of music by comparing the effects of daily vocal music, instrumental music, and audio book listening using a randomized controlled trial design.

Compared to instrumental music or audiobooks, vocal music listening was found to enhance the recovery of verbal memory across patients and language skills specifically in aphasic patients, indicating that the singing component is crucial for the rehabilitative effect of music listening. From a conceptual point of view, vocal songs represent an interface between speech and music. They bind the lyrics and melody into a unified representation and provide a structured temporal scaffolding framework that facilitates the recall (Ferreri and Verga, 2016). In general, the close coupling between vocal music and verbal memory is evidenced by studies showing that sung melody enhances the learning and recall of novel verbal material in healthy subjects (Wallace, 1994; Kilgour et al., 2000), stroke patients (see Studies I and II), and patients with Alzheimer's disease (Moussard et al., 2014; Palisson et al., 2015). Previous studies have also revealed that overt production of verbal material during memory encoding is more effective for later recall when it is done through singing than through speaking, both in healthy subjects (Quinlan and Taylor, 2013) and in aphasic stroke patients (Racette et al., 2006). Singing-based interventions have also been shown to enhance verbal memory in healthy older adults (Degé and Kerkovius, 2018) and in persons with mild dementia (Särkämö et al., 2016; Pongan et al., 2017).

The close linkage between vocal music and verbal working memory is conversely evidenced by a larger disruptive effect of vocal compared to instrumental background music in a verbal working memory task reported in healthy subjects (Salamé and Baddeley, 1989). This effect has been attributed to the fact that vocal music has more or less automatic access to the phonological loop of working memory where it affects subvocal rehearsal. In Study III, the patients in the VMG were not instructed to sing along to the songs; however, it is plausible that listening to songs with lyrics may have elicited some subvocal processing which could have covertly trained working memory, over time leading to the enhanced recovery seen in the verbal memory tasks.

Also, the enhancement of language skills induced by the vocal music listening may be related to the subvocal training since singing-based rehabilitation has been found to be effective for speech recovery in aphasia (van der Meulen et al., 2014; Raglio et al., 2016).

In addition to engaging verbal working memory, vocal music evokes increased vigilance or arousal, resulting in greater depth of processing which enhances memory for sung material (Weiss, Trehub, & Schellenberg, 2012). This effect is likely mediated by emotional factors and the reward value of music. Compared to speech, music is more closely associated with pleasant and rewarding experiences which arise from increased dopaminergic activation of the mesolimbic reward system comprising multiple limbic areas, such as the ventral striatum, amygdala, and hippocampus, as well as medial frontal areas, including the anterior cingulate, orbitofrontal cortex, and medial prefrontal cortex (Koelsch, 2014; Salimpoor et al., 2013). It has been shown that administration of a dopaminergic precursor (levodopa) enhances verbal learning and the experienced pleasure of the learning process (Ripollés et al., 2018). Reported pleasure and measured engagement of the limbic network are higher when listening to familiar versus unfamiliar music (Pereira et al., 2011; van den Bosch et al., 2013) and sung versus instrumental music (Brattico et al., 2011; Alluri et al., 2013). Given this and the observation that the individual reward value of music mediates its positive effect on episodic memory (Ferreri et al., 2017), it is possible that the rehabilitative effect of vocal music on verbal memory in Study III is at least partly driven by its intrinsic ability to engage motivation- and reward-related dopaminergic networks. In converge with this, a recent study showed that pleasure experienced during music listening is mediated by dopaminergic signaling (Ferreri et al., 2019). In general, music listening has a robust mood-enhancing and stress-reducing effect in daily life (Linnemann, Ditzen, Strahler, Doerr & Nater, 2015; Finn and Fancourt, 2018), which can also explain the lower POMS Confusion and Depression scores in the music group than in the ABG.

Previously, a larger increase in grey matter volume in stroke patients with left hemisphere damage was reported in left and right SFG, left anterior cingulate, and right ventral striatum after daily music listening compared to audiobooks and standard care (Särkämö et al., 2014). Using a larger sample (N = 75) and more rigorous statistical criteria (FWE correction), the voxel-based morphometry results of Study III showed that compared to audiobooks, music listening specifically increased grey matter volume in left anterior temporal (STG, MTG, ITG) and parietal (IPL, AG, precuneus, MOG) areas across all patients, with effects in the left temporal cluster seen especially for vocal music. In addition, previous fMRI studies of healthy subjects have reported stronger activation in left anterior temporal regions during listening to songs vs. speech (Callan et al., 2006; Sammler et al., 2010; Schön et al., 2010; Whitehead and Armony, 2018). The left IPL and surrounding posterior temporal and parietal regions, in turn, play a crucial role in perceiving the melodic and rhythmic structure of music (Lee, Janata, Frost, Hanke & Granger, 2011; Konoike et al., 2012) as well as in sensory-motor integration of speech (Hickok and Poeppel, 2007) and verbal working memory, especially in subvocal rehearsal (Paulesu, Frith & Frackowiak, 1993; Fegen, Buchsbaum & D'Esposito, 2015). Notably, the MG > ABG and VMG > ABG effects were partly driven by a reduction of grey matter volume in the ABG in these clusters. Previous studies suggest that both increases and decreases in grey matter volume take place after stroke. Although the precise cellular and molecular mechanism underlying grey matter volume increase are still largely unknown, it has been suggested to result from increases in axon sprouting, dendritic branching, and glial changes as well as in the production or formation of new neurons (neurogenesis), synapses (synaptogenesis), and blood vessels (angiogenesis), resulting in learning- and recovery-enhancing neuroplasticity (Zatorre, Fields & Johansen-Berg, 2012). Grey matter volume decrease, in turn, represents atrophy and is associated with poorer clinical recovery (Stebbins et al., 2008; Yang et al., 2019) and lower gains induced by environmental enrichment (Mattsson, Sorenson, Zimmer & Johansson, 1997; Hase, O'Brien, Moore & Freeman, 2018) or rehabilitation (Gauthier, Taub, Mark, Barghi & Uswatte 2012; Moore et al., 2015). It is possible that the large-scale activation of cortical and subcortical areas induced by music listening after stroke (Sihvonen et al., 2017b) can have a long-term neuroprotective impact by preventing atrophy in cortical areas most strongly activated by songs.

The voxel-based morphometry results within the aphasic subgroup revealed a somewhat different pattern of volume changes induced by music listening. In aphasic patients, music listening increased grey matter volume compared to audiobooks in the right precuneus and angular gyrus (AG), two key structures for experiencing music-evoked emotions (Blood & Zatorre, 2001; Tabei, 2015; Rogenmoser, Zollinger, Elmer & Jäncke, 2016) and episodic memories (Platel, Baron, Desgranges, Bernard & Eustache, 2003; Altenmüller et al., 2014; Sikka, Cuddy, Johnsrude & Vanstone, 2015), with additional volume increases induced by vocal / instrumental music in neighbouring posterior temporal (MTG) and parieto-occipital (LG, cuneus, MOG) areas, which have been linked to music and speech perception (Angulo-Perkins et al., 2014; Hymers et al., 2015) and memory-related visual imagery (Gardini, Cornoldi, De Beni & Venneri 2006; Bastin et al., 2013). Importantly, the volume changes in these posterior temporal and parietal regions correlated with enhanced music-induced recovery of language and verbal memory. This is in line with previous studies showing language reorganization after left frontotemporal damage (Kinno et al., 2014; Yang et al., 2018) and therapy-induced changes in aphasia (Musso et al., 1999; Fridriksson et al. 2007; Abel, Weiller, Huber, Willmes & Specht, 2015) in these regions.

In addition to the structural changes, there were also long-term music-induced functional connectivity changes in the default mode network. Being functionally linked to emotional processing, self-referential mental activity, and the recollection of prior experiences (Raichle, 2015; Belfi et al., 2018), the default mode network is strongly engaged during music listening, especially when the music is familiar and pleasant (Wilkins, Hodges, Laurienti, Steen & Burdette, 2014; Taruffi, Pehrs, Skouras & Koelsch, 2017), also in non-musician listeners (Alluri et al., 2017). In line with this, the VMG showed larger increase in functional connectivity in the whole default mode network during vocal music (but not instrumental music or speech) listening than the ABG, indicating functional neuroplasticity specific to the type of stimulus and intervention. Importantly, also resting-state functional connectivity in the left temporal (STG/MTG) areas of the default mode network increased more in VMG than ABG and IMG, and correlated with the improved recovery of language and verbal memory. This finding is supported by previous studies showing that reduced default mode network connectivity is associated with the impairment of verbal memory in ageing (Staffaroni et al., 2018; Zhang et al., 2019) or after stroke (Tuladhar et al., 2013). Good longitudinal recovery after stroke is, in turn, linked to increased default mode network connectivity (Dacosta-Aguayo et al., 2014; Park et al., 2014). Together, the voxel-based morphometry and functional connectivity results provide compelling evidence that the rehabilitative effect of vocal music is underpinned by both structural and functional plasticity changes in temporoparietal networks that are crucial for emotional processing, language, and memory.

Study III has some methodological limitations, which should be considered when evaluating its findings. First of al, the study was a pooled analysis of two randomized controlled trials, one with randomization to three (VMG/IMG/ABG) groups (Turku) and the other with randomization to two groups (MG/ABG) and then post hoc reclassification to the VMG and ABG (Helsinki). It is therefore possible that there is a slight self-selection bias, more so between the two music groups and less so between the music groups and the ABG. However, since the study design of the two trials was otherwise highly similar and that all outcome measure results were covaried for trial site, the bias is unlikely significant. Second, there were relatively small sample sizes in all three studies in the subgroup analyses of aphasic patients (N=29) and in the fMRI analyses (N=35). Therefore, the results should be considered somewhat tentative and need to be confirmed with larger studies.

While more research is still needed, the natural future steps could be to uncover in a larger trial the impact of the different demographic and clinical background factors to the efficacy of music listening, such as pre-stroke music listening or amusia. In addition, larger trials of aphasic patients with more aphasia subtypes and varying severity levels of aphasia could lead the way towards better understanding of the therapeutic role and mechanisms of music listening in aphasia, potentially leading to more individualized use of music in aphasia rehabilitation. In future research, it would also be interesting and important to compare the effects of more active singing-based rehabilitation and more passive listening-based music rehabilitation. Specifically, considering the findings of Studies I and II, one future direction could also be to determine if sung memory recall, or more active (overt) rehearsal of the verbal material through singing, could be used to enhance learning even further and also to have a more general transfer effect on working memory or verbal memory for untrained material after stroke and particularly in aphasia. The same rehabilitation approach could be studied also in other patient groups suffering from verbal memory deficits, such as patients with multiple sclerosis.

5.4 Conclusions and clinical considerations

The present thesis demonstrates the pre-eminent role of the vocal music as an effective tool for facilitating verbal learning and for enhancing verbal, cognitive, emotional, and neural recovery after stroke. The latter finding is supported by the previous studies that have found a positive effect of daily music listening on stroke recovery (Särkämö et al., 2008, 2010, 2014; Baylan et al. 2018, 2019). However, the findings of this thesis show for the first time that the sung component of music is a key factor driving the rehabilitative effect of music listening on verbal memory. Crucially, the novel results also show that vocal music listening can speed up language recovery in aphasia during the first three months, likely by providing additional stimulation to bilateral frontotemporal networks that are normally upregulated in language recovery during this stage (Hartwigsen & Saur, 2017).

Clinically, the findings address a vital issue of how the patient environment can be optimized for recovery during the first weeks after stroke when typically over 70% of daily time is spent in non-therapeutic activities (Bernhardt et al. 2004; De Wit et al., 2005) even though this time-window would be ideal for rehabilitation from the standpoint of brain plasticity. Corroborating previous findings (Särkämö et al., 2008, 2010, 2014; Baylan et al. 2018, 2019), the present thesis provides further evidence for the use of daily music listening, especially of vocal music, as an effective, easily applicable, and inexpensive way to support cognitive and emotional recovery after stroke. Importantly, music listening can be used both to optimize the recovery time at the hospital, for example as a part of broader environmental enrichment intervention (Janssen et al., 2014; Khan et al., 2016; Rosbergen et al., 2017) and as a pleasant and motivating leisure activity at home. Particularly the finding that vocal music can facilitate language recovery at the early post-stroke stage is clinically important as it suggests that vocal music listening could be used to supplement speech therapy for aphasia, which is often difficult to implement during the first post-stroke weeks and months due to the severity of symptoms, general fatigue, and lack of rehabilitation resources.

6 REFERENCES

Abel, S., Weiller, C., Huber, W., Willmes, K. & Specht, K. Therapy-induced brain reorganization patterns in aphasia. *Brain, 138*, 1097–1112 (2015).

Aivoliitto. Aivoverenkiertohäiriö Suomessa. Retrieved November 24, 2019 from https://dyajetwym1cg9.cloudfront.net › files › avh_lukuina2013_web (2013).

Albert, M. L., Sparks, R. W., & Helm, N. A. Melodic intonation therapy for aphasia. *Arch Neurol, 29*, 130-131(1973).

Alluri, V., Toiviainen, P., Jääskeläinen, I. P., Glerean, E., Sams, M. & Brattico, E. Large-scale brain networks emerge from dynamic processing of musical timbre, key and rhythm. *Neuroimage, 59*, 3677-89 (2012).

Alluri, V., Toiviainen, P., Lund, T. E., Wallentin, M., Vuust, P., Nandi, A. K. Ristaniemi, T. & Brattico, E. From Vivaldi to Beatles and back: predicting lateralized brain responses to music. *Neuroimage, 83*, 627-36 (2013).

Altenmüller, E., Marco-Pallares, J., Münte, T. F. & Schneider, S. Neural reorganization underlies improvement in stroke-induced motor dysfunction by music-supported therapy. *Ann N Y Acad Sci, 1169,* 395-405 (2009).

Anderson, S., Himondes, E., Wise, K., Welch, G. & Stewart, L. Is there potential for learning in amusia? A study of the effect of singing intervention in congenital amusia. *Ann N Y Acad Sci, 1252*, 345-53 (2012).

Angulo-Perkins, A., Aubé, W., Peretz, I., Barrios, F. A., Armony, J. L. & Concha, L. Music listening engages specific cortical regions within the temporal lobes: differences between musicians and non-musicians. *Cortex, 59*, 126–137 (2014).

Ashburner, J., & Friston, K. J. Voxel-based morphometry--the methods. *Neuroimage, 11*, 805-821 (2000).

Ashburner, J. & Friston, K. J. Unified segmentation. *Neuroimage, 26*, 839-51 (2005).

Atkinson, R. C. & Shiffrin, R. M. Human memory: A proposed system and its control processes. *Psychol Learn Motiv, 2*, 89-195 (1968).

Ayotte, J., Peretz, I., Rousseau I., Bard, C. & Bojanowski, M. Patterns of music agnosia associated with middle cerebral artery infarcts. *Brain, 123*, 1926-1938 (2000).

Barnes, J., Ridgway, G. R., Bartlett, J., Henley, S. M., Lehmann, M., Hobbs, N., Clarkson, M. J., MacManus, D. G., Ourselin, S. & Fox, N. C. Head size, age and gender adjustment in MRI studies: a necessary nuisance? *Neuroimage, 53*, 1244–1255 (2010).

Bastin, C., Feyers, D., Jedidi, H., Bahri, M. A., Degueldre, C., Lemaire, C., Collette, F. & Salmon E. Episodic autobiographical memory in amnestic mild cognitive impairment: what are the neural correlates? *Hum Brain Mapp*,34, 1811-25 (2013).

Baylan, S., Haig, C., MacDonald, M., Stiles, C., Easto, J., Thomson, M., Cullen, B., Quinn, T. J., Stott, D., Mercer, S. W., Broomfield, N. M., Murray, H. & Evans, J. J. Measuring the effects of listening for leisure on outcome after stroke (MELLO): A pilot randomized controlled trial of mindful music listening. *Int J Stroke*, Apr 2:1747493019841250 (2019).

Baylan, S., McGinlay, M., MacDonald, M., Easto, J., Cullen, B., Haig, C., Mercer, S. W., Murray, H., Quinn, T. J., Stott, D., Broomfield, N. M., Stiles, C. & Evans, J. J. Participants' experiences of music, mindful music, and audiobook listening interventions for people recovering from stroke. *Ann N Y Acad Sci, 1423*, 349-359 (2018).

BDI-II manual. Beck, A. T., Steer, R. A., & Brown, G. K. San Antonio, TX: The Psychological Corporation (1996).

Belfi, A. M., Vessel, E. A., Brielmann, A., Isik, A. I., Chatterjee, A., Leder, H., Pelli, D. G. & Starr, G. G. Dynamics of aesthetic experience are reflected in the default-mode network. *Neuroimage*, 188, 584-97 (2019).

Belin, P., Van Eeckhout, P., Zilbovicius, M., Remy, P., François, C., Guillaume, S., Chain, F., Rancurel, G. & Samson, Y. Recovery from nonfluent aphasia after melodic intonation therapy: a PET study. *Neurology, 47*, 1504-1511 (1996).

Bell, A. J. & Sejnowski, T. J. An information-maximization approach to blind separation and blind deconvolution. *Neural Comput,* 7, 1129-59 (1995).

Bernhardt, J., Dewey, H., Thrift, A. & Donnan, G. Inactive and alone. Physical activity within the first 14 days of acute stroke unit care. *Stroke, 35*, 1005–1009 (2004).

Blood, A. J. & Zatorre, R. J. Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proc Natl Acad Sci U S A, 98*, 11818–11823 (2001).

Bonakdarpour, B., Eftekharzadeh, A. & Ashayeri, H. Melodic Intonation Therapy in Persian aphasic patients. *Aphasiology, 17*, 75-95.

Bor, D., Cumming, N., Scott, C. E. & Owen, A. M. Prefrontal cortical involvement in verbal encoding strategies. *Eur J Neurosci, 19*, 3365-70 (2004).

Bor, D., Duncan, J., Wiseman, R. J., & Owen, A. M. Encoding strategies dissociate prefrontal activity from working memory demand. *Neuron, 37*, 361-367 (2003).

Bosseler, A. N., Teinonen, T., Tervaniemi, M., & Huotilainen, M. Infant-directed speech enhances statistical learning in newborn infants: An ERP study. *PLoS One, 11*, e0162177 (2016).

Brattico, E., Alluri, V., Bogert, B., Jacobsen, T., Vartiainen, N., Nieminen, S. & Tervaniemi, M. A functional MRI study of happy and sad emotions in music with and without lyrics. *Front Psychol, 2*, 308 (2011).

Breier, J. I., Randle, S., Maher, L. M., & Papanicolaou, A. C. Changes in maps of language activity activation following melodic intonation therapy using magnetoencephalography: two case studies. *J Clin Exp Neuropsychol, 32*, 309-314 (2010).

Brett, M., Leff, A. P., Rorden, C. & Ashburner, J. Spatial normalization of brain images with focal lesions using cost function masking. *Neuroimage*, *14*, 486-500 (2001).

Brown, S. A joint prosodic origin of language and music. *Front Psychol, 8*, 1894, (2017).

Brown, S., Martinez, M. J., Hodges, D. A., Fox, P. T. & Parsons, L. M. The song system of the human brain. *Brain Res, 20*, 363-375 (2004).

Buckner, R. L., Andrews-Hanna, J. R. & Schacter, D. L. The Brain's Default Network: Anatomy, Function, and Relevance to Disease. *Ann N Y Acad Sci, 1124,* 1–38 (2008).

Buchsbaum, B. R., Olsen, R. K., Koch, P. & Berman, K. F. Human dorsal and ventral auditory streams subserve rehearsal-based and echoic processes during verbal working memory, *Neuron, 48*, 687-697 (2005).

Buchsbaum, B. R., Ye, D., & D'Esposito, M. Recency effects in the inferior parietal lobe during verbal recognition memory. *Front Hum Neurosci, 5*, 59 (2011).

Callan, D. E., Tsytsarev, V., Hanakawa, T., Callan, A. M., Katsuhara, M., Fukuyama, H. & Turner, R. Song and speech: brain regions involved with perception and covert production. *Neuroimage*, 31, 1327-42 (2006).

Calvert, S. L., & Tart, M. Song versus verbal forms for very-long-term, long-term, and short-term verbatim recall. *J Appl Dev Psychol*, *14*, 245-260 (1993).

Chazin, S. & Neuschatz, J. S. Using a mnemonic to aid in the recall of unfamiliar information. *Percept Mot Ski, 71***,** 1067–1071 (1990).

Chiou, K. S., Genova, H. M., & Chiaravalloti, N. D. Structural white matter differences underlying heterogeneous learning abilities after TBI. *Brain Imaging Behav, 10*, 1274-1279 (2016).

Colby, J. B., Soderberg, L., Lebel, C., Dinov, I. D., Thompson, P. M., & Sowell, E. R. Along-tract statistics allow for enhanced tractography analysis. *Neuroimage, 59*, 3227-3242 (2012).

Cramer, S. C. Repairing the human brain after stroke: I. Mechanisms of spontaneous recovery. *Ann Neurol, 63*, 272-87 (2008).

Crinion, J., Ashburner, J., Leffa, A., Brett, M., Price, C. & Friston, K. Spatial normalization of lesioned brains: Performance evaluation and impact on fMRI analyses. *Neuroimage, 37*, 866-875 (2007).

Crowder, R.G., Serafine, M. L. & Repp, B.H. Physical interaction and association by contiguity in memory for the words and melodies of songs. *Mem Cognit, 18*, 469–476 (1990).

Dacosta-Aguayo, R., Graña, M., Savio, A., Fernández-Andújar, M., Millán, M., López-Cancio, E., Cáceres, C., Bargalló, N., Garrido, C., Barrios, M., Clemente, I. C., Hernández, M., Munuera, J., Dávalos, A., Auer, T. & Mataró, M. Prognostic value of changes in resting-state functional connectivity patterns in cognitive recovery after stroke: a 3T fMRI pilot study. *Hum Brain Mapp, 35*, 3819–3831 (2014).

Dahlqvist, P., Rönnbäck, A., Bergström, S. A., Söderström, I. & Olsson, T. Environmental enrichment reverses learning impairment in the Morris water maze after focal cerebral ischemia in rats. *European J Neurosci, 19*, 2288-98 (2004).

Deese, J., & Kaufman, R. A. Serial effects in recall of unorganized and sequentially organized verbal material. *J Exp Psychol, 54*, 180-187 (1957).

Degé, F. & Kerkovius, K. The effects of drumming on working memory in older adults. *Ann N Y Acad Sci, 423*, 242-250 (2018).

De Renzi, E. & Faglioni, P. Normative data and screening power of a shortened version of the Token Test. *Cortex,* 14, 41–49 (1978).

De Wit, L., Putman, K., Dejaeger, E., Baert, I., Berman, P., Bogaerts, K., Brinkmann, N., Connell, L., Feys, H., Jenni, W., Kaske, C., Lesaffre, E., Leys, M., Lincoln, N., Louckx, F., Schuback, B., Schupp, W., Smith, B. & Weerdt, W. Use of time by stroke patients: a comparison of four European rehabilitation centers. *Stroke, 36*, 1977-83 (2005).

Dick, A. S. & Tremblay, P. Beyond the arcuate fasciculus: Consensus and controversy in the connectional anatomy of language. *Brain, 135*, 3529-50 (2012).

DiLuca, M. & Olesen, J. The cost of brain diseases: a burden or a challenge? *Neuron, 82*, 1205-8 (2014).

Dobkin, B.H. Strategies for stroke rehabilitation. *The Lancet Neurol 3*, 528-36 (2004).

Duffau, H., Herbet, G. & Moritz-Gasser, S. Toward a pluri-component, multimodal, and dynamic organization of the ventral semantic stream in humans: lessons from stimulation mapping in awake patients. *Front. Syst. Neurosci, 7*, 1-4 (2013).

Düzel, E., Hufnagel, A., Helmstaedter, C., & Elger, C. Verbal working memory components can be selectively influenced by transcranial magnetic stimulation in patients with left temporal lobe epilepsy. *Neuropsychologia, 34*, 775-783 (1996).

Edits manual for the Profile of Mood States. McNair, D. M., Lorr, M. & Droppleman, L. F. California: Educational and industrial testing services (1981).

Engelter, S. T., Gostynski, M., Papa, S., Frei, M., Born, C., Ajdacic-Gross, Gutzwiller, F. & Lyrer, P. A. Epidemiology of aphasia attributable to first ischemic stroke. Incidence, severity, fluency, etiology, and thrombolysis. *Stroke, 37*, 1379–1384 (2006).

Eom, B., & Sung, J. E. The effects of sentence repetition-based working memory treatment on sentence comprehension abilities in individuals with aphasia. *Am J Speech Lang Pathol, 25*, S823-838 (2016).

Fedorenko, E., Patel, A., Casasanto, D., Winawer, J. & Gibson, E. (2009). Structural integration in language and music: evidence for a shared system. *Mem Cognit*, *37*, 1–9.

Fegen, D., Buchsbaum, B. R. & D'Esposito, M. The effect of rehearsal rate and memory load on verbal working memory. *Neuroimage, 105*, 120-131 (2015).

Feigin, V. L., Forouzanfar, M. H., Krishnamurthi, R., Mensah, G. A., Connor, M., Bennett, D. A., Moran, A. E., Sacco, R. L., Anderson, L., Truelsen, T., O'Donnell, M., Venketasubramanian, N., Barker-Collo, S., Lawes, C. M., Wang, W., Shinohara, Y., Witt, E., Ezzati, M., Naghavi, M. & Murray, C.; Global burden of diseases, injuries, and risk factors study 2010 (GBD 2010) and the GBD stroke experts group. Global and regional burden of stroke during 1990-2010: findings from the global burden of disease study. *Lancet, 383*, 245-54 (2010).

Ferreri, L., Bigand, E., Bard, P., & Bugaiska, A. The influence of music on prefrontal cortex during episodic encoding and retrieval of verbal information: A multichannel fNIRS study. *Behav. Neurol., 2015*, article ID 707625 (2015).

Ferreri, L. & Verga, L. Benefits of music on verbal learning and memory: how and when does it work? *Music Percept, 34*, 167-82 (2016).

Flowers, H. L., Flowers, F. L., Silver, J. Fang, E., Rochon, E. & Martino, R. The incidence, co-occurrence and predictors of dysphagia, dysarthria, and aphasia after first-ever acute ischemic stroke. *J Commun Disord, 46*, 238–248 (2013).

Forsblom, A., Laitinen, S., Särkämö, T. & Tervaniemi, M. Therapeutic role of music listening in stroke rehabilitation. *Ann N Y Acad Sci, 1169*, 426-30 (2009).

Fridriksson, J., Moser, D., Bonilha, L., Morrow-Odom, K. L., Shaw, H., Fridriksson, A., Baylis, G. C. & Rorden, C. Neural correlates of phonological and semantic-based anomia treatment in aphasia. *Neuropsychologia*, *45*, 1812-22 (2007).

Gardini, S., Cornoldi, C., De Beni, R. & Venneri, A. Left mediotemporal structures mediate the retrieval of episodic autobiographical mental images. *Neuroimage, 30*, 645–655 (2006).

Gauthier, L. V., Taub, E., Mark, V. W., Barghi, A. & Uswatte, G. Atrophy of spared gray matter tissue predicts poorer motor recovery and rehabilitation response in chronic stroke. *Stroke, 43*, 453–457 (2012).

GBD 2016 Stroke Collaborators. Global, regional, and national burden of stroke, 1990-2016: a systematic analysis for the global burden of disease study 2016. *Lancet Neurol, 18*, 439–458 (2019).

Global Health Estimates. Geneva: World Health Organization. Retrieved November 24, 2019 from: http://www.who.int/healthinfo/global_burden_disease/en/ (2012).

Gobbo, O. L. & O'Mara, S. M. Impact of enriched-environment housing on brain-derived neurotrophic factor and on cognitive performance after a transient global ischemia. *Behav Brain Res, 152*, 231-241, (2004).

Gobet, F., Lane, P. C., Croker, S., Cheng, P. C., Jones, G., Oliver, I. & Pine, J. M. Chunking mechanisms in human learning. *Trends Cogn Sci, 5*, 236-243 (2001).

Grahn, J. A. & Rowe, J. B. Feeling the beat: Premotor and striatal interactions in musicians and nonmusicians during beat perception. *Neurosci, 29,* 7540-7548 (2009).

Grau-Sánchez, J., Duarte, E., Ramos-Escobar, N., Sierpowska, J., Rueda, N., Redón, S., Veciana de Las Heras, M., Pedro, J., Särkämö, T. & Rodríguez-Fornells, A. Music-supported therapy in the rehabilitation of subacute stroke patients: a randomized controlled trial. *Ann N Y Acad Sci*, *1423*, 318-28 (2018).

Group ICA Of fMRI Toolbox (GIFT) http://mialab.mrn.org/software/gift/

Halwani, G. F., Loui, P., Rüber, G. & Schlaug, G. Effects of practice and experience on the arcuate fasciculus: comparing singers, instrumentalists, and non-musicians. *Front. Psychol., 2*, 156 (2011).

Hartwigsen, G. & Saur, D. Neuroimaging of stroke recovery from aphasia - insights into plasticity of the human language network. *Neuroimage*, *190*, 14-31 (2017).

Hase, A., O'Brien, J., Moore, L. J. & Freeman, P. The relationship between challenge and threat states and performance: A systematic review. *Sport Exerc Perform Psychol, 8*, 123-144 (2018).

Haslam, C., & Cook, M. Striking a chord with amnestic patients: Evidence that song facilitates memory. *Neurocase, 8,* 453-465 (2002).

Hébert, S., Racette, A., Gagnon, L., & Peretz, I. Revisiting the dissociation between singing and speaking in expressive aphasia. *Brain*, *126*, 1838-1850 (2003).

Herholz, S. C. & Zatorre, R. J. Musical training as a framework for brain plasticity: behavior, function, and structure. *Neuron, 76*, 486-502 (2012).

Hickok, G. & Poeppel, D. The cortical organization of speech processing. *Nat Rev Neurosci, 8*, 393–402 (2007).

Hilari, K., Byng, S., Lamping, D. L. & Smith, S. C. Stroke and aphasia quality of life scale-39 (SAQOL-39): evaluation of acceptability, reliability, and validity. *Stroke, 34,* 1944-50 (2003).

Hänninen, H. Neurotoksisten haittojen seulonta - oirekyselyt ja psykologiset testit. Helsinki, Finland: Työterveyslaitos (1989). In Finnish.

Hymers, M., Prendergast, G., Liu, C., Schulze, A., Young, M. L., Wastling, S. J., Barker, G. J. & Millman, R. E. Neural mechanisms underlying song and speech perception can be differentiated using an illusory percept. *Neuroimage*, 108, 225-33 (2015).

Inatomi, Y., Yonehara, T., Omiya, S., Hashimoto, Y., Hirano, T. & Uchino, M. Aphasia during the acute phase in ischemic stroke. *Cerebrovasc Dis, 25,* 316–323 (2008).

Innocenti, I., Cappa, S. F., Feurra, M., Giovannelli, F., Santarnecchi, E., Bianco, G., Cincotta, M. & Rossi, S. TMS interference with primacy and recency mechanisms reveals bimodal episodic encoding in the human brain. *J Cogn Neurosci, 25,* 109-116 (2013).

Isarida, T. & Isarida, T. K. Influences of environmental context on the recency effect in free recall. *Mem Cognit, 34,* 787-794 (2006).

Ivanova, M. V., Dragoy, O., Kuptsova, S. V., Yu Akinina, S., Petrushevskii, A. G., Fedina, O. N.,Turken, A., Shklovsky, V. M. & Dronkers, N. F. Neural mechanisms of two different verbal working memory tasks: A VLSM study. *Neuropsychologia, 115*, 25-41 (2018).

Janata, P. The neural architecture of music-evoked autobiographical memories. *Cereb Cortex, 19*, 2579-2594 (2009).

Janata, P., Birk, J. L., Van Horn, J. D., Leman, M., Tillmann, B., & Bharucha, J. J. The cortical topography of tonal structures underlying Western music. *Science, 298,* 2167-2170 (2002).

Janssen, H., Ada, L., Bernhardt, J., McElduff, P., Pollack, M., Nilsson, M. & Spratt, N. J. An enriched environment increases activity in stroke patients undergoing rehabilitation in a mixed rehabilitation unit: a pilot non-randomized controlled trial. *Disabil Rehabil, 36*, 255-62 (2014).

Jefferies, E., Hoffman, P., Jones, R., & Ralph, M. A. The impact of semantic impairment on verbal short-term memory in stroke aphasia and semantic dementia: A comparative study. *J Mem Lang, 58,* 66-87 (2008).

Jerde, T. A., Childs, S. K., Handy, S. T., Nagode, J. C., & Pardo, J. V. Dissociable systems of working memory for rhythm and melody. *Neuroimage, 57*, 1572-1579 (2011).

Johansson, B. B. Functional and cellular effects of environmental enrichment after experimental brain infarcts. *Restor Neurol Neurosci, 22*, 163-74 (2004).

Jungblut, M., Huber, W., Mais, C., & Schnitker, R. Paving the way for speech: voice-training-induced plasticity in chronic aphasia and apraxia of speech--three single cases. *Neural Plast, 2014*, 841982 (2014).

Kadojic, D., Bijelic, B.R., Radanovic, R., Porobić, M., Rimac, J. & Dikanović, M. Aphasia in patients with ischemic stroke. *Acta Clin Croat, 51*, 221-225 (2012).

Kasdan, A., & Kiran, S. Please don't stop the music: Song completion in patients with aphasia. *J Commun Disord, 75*, 72-86 (2018).

Khan, F., Amatya, B., Elmalik, A., Lowe, M., Ng, L., Reid, I. & Galea, M. P. An enriched environmental programme during inpatient neuro-rehabilitation: a randomized controlled trial. *J Rehabil Med, 48*, 417-25 (2016).

Kilgour, A. R., Jakobson, L. S., & Cuddy, L. L. Music training and rate of presentation as mediators of text and song recall. *Mem Cognit, 28*, 700-710 (2000).

Kinno, R., Ohta, S., Muragaki, Y., Maruyama, T. & Sakai, K. L. Differential reorganization of three syntax-related networks induced by a left frontal glioma. *Brain*, *137*, 1193-212 (2014).

Kleber, B., Veit, R., Birbaumer, N., Gruzelier, J. & Lotze, M. The Brain of opera singers: experience-dependent changes in functional activation. *Cereb Cortex, 20*, 1144–1152 (2010).

Koelsch, S. Brain correlates of music-evoked emotions. *Nat Rev Neurosci,* 15, 170-80 (2014).

Konoike, N., Kotozaki, Y., Miyachi, C. M. Yomogida, Y., Akimoto, Y., Kuraoka, K., Sugiura, M., Kawashima, R. & Nakamura, K. Rhythm information represented in the fronto-parieto-cerebellar motor system. *Neuroimage, 63*, 328-38 (2012).

Krogsrud, S. K., Fjell, A. M., Tamnes, C. K., Grydeland, H., Due-Tønnessen, P., Bjørnerud, A., Sampaio-Baptista, C., Andersson, J., Johansen-Berg, H. & Walvhovd, K. B. Development of white matter microstructure in relation to verbal and visuospatial working memory – A longitudinal study. *PLoS One, 13,* e0195540 (2018).

LaCroix, A. N., Diaz, A. F. & Rogalsky, C. The relationship between the neural computations for speech and music perception is context-dependent: an activation likelihood estimate study. *Front Psychol, 6*, 1138 (2015).

Laine, M., Goodglass, H., Niemi, J., Koivuselkä-Sallinen, P., Tuomainen, J., & Marttila, R. Adaptation of the Boston Diagnostic Aphasia Examination and the Boston Naming Test into Finnish. *Scand J Log Phon*, *18*, 83-92 (1993a).

Laine, M., Tuomainen, J., & Ahonen, A. Changes in hemispheric brain perfusion elicited by melodic intonation therapy: A preliminary experiment with single photon emission computed tomography (SPECT). *Scand J Log Phon, 19*, 19-24 (1993b).

Laird, A. R., Eickhoff, S. B., Li, K., Robin, D. A., Glahn, D. C. & Fox, P. T. Investigating the functional heterogeneity of the default mode network using coordinate-based meta-analytic modeling. *J Neurosci, 29*, 14496-14505 (2009).

Langhorne, P., Bernhardt, J. & Kwakkel, G. Stroke rehabilitation. *Lancet 377*, 1693-1702 (2011).

Lee, Y. S., Janata, P., Frost, C., Hanke, M. & Granger, R. Investigation of melodic contour processing in the brain using multivariate pattern-based fMRI. *Neuroimage, 57*, 293–300 (2011).

Leemans, A. & Jones, D. K. The B-matrix must be rotated when correcting for subject motion in DTI data. *Magn Reson Med, 61*, 1336-1349 (2009).

Leroy, F., Glasel, H., Dubois, J., Hertz-Pannier, L., Thirion, B., Mangin, J. & Dehane-Lambertz, G. Early maturation of the linguistic dorsal pathway in human infants, *J Neurosci, 31*, 1500-1506 (2011).

Lévêque, Y. & Schön, D. Modulation of the motor cortex during singing-voice perception. *Neuropsychologia*, *70*, 58-63 (2015).

Lieberman, M. D. & Cunningham, W. A. Type I and type II error concerns in fMRI research: Re-balancing the scale. *Soc Cogn Affect Neurosci, 4,* 423-428 (2009).

Livingston-Thomas, J., Nelson, P., Karthikeyan, S., Antonescu, S., Jeffers, M., Marzolini, S. & Corbett, D. Exercise and environmental enrichment as enablers of task-specific neuroplasticity and stroke recovery. *Neurotherapeutics, 13*, 305–402 (2016).

Linnemann, A., Ditzen, B., Strahler, J., Doerr, J. M. & Nater, U. M. Music listening as a means of stress reduction in daily life. *Psychoneuroendocrinology, 60*, 82–90 (2015).

Loui, P. A dual-stream neuroanatomy of singing. *Music percept, 32*, 232-241 (2015).

López-Barroso, D., Catani, M., Ripollés, P., Dell'Acqua, F., Rodríguez-Fornells, A., & de Diego-Balaguer, R. Word learning is mediated by the left arcuate fasciculus. *Proc Natl Acad Sci U S A, 110*, 13168-13173 (2013).

López-Barroso, D., Ripollés, P., Marco-Pallarés, J., Mohammadi, B., Münte, T. F., Bachoud-Lévi, A. C., Rodriguez-Fornells, A. & de Diego-Balaguer, R. Multiple brain networks underpinning word learning from fluent speech revealed by independent component analysis. *Neuroimage*, *110*, 182-93 (2015).

Ludke, K. M., Ferreira, F., & Overy, K. Singing can facilitate foreign language learning. *Mem Cognit 42*, 41-52 (2014).

Mabbott, D. J., Rovet, J., Noseworthy, M. D., Smith, M. L., & Rockel, C. The relations between white matter and declarative memory in older children and adolescents. *Brain Res., 1294*, 80-90 (2009).

Mackey, F., Ada, L., Heard, R. & Adams, R. Stroke rehabilitation: Are highly structured units more conducive to physical activity than less structured units? *Arch Phys Med Rehabil, 77*, 1066-1070 (1996).

Magee, W. L., Clark, I., Tamplin, J. & Bradt, J. (2017). Music interventions for acquired brain injury. *Cochrane Database Syst Rev*, *20*, 1.

Manual of aphasia and aphasia therapy 2nd ed. Helm-Estabrooks, N. & Albert, M. L. Pro-Ed; Austin, TX, (2004) 221–233.

Mas-Herrero, E., Marco-Pallares, J., Lorenzo-Seva, U., Zatorre, R. J., & Rodriguez-Fornells, A. Individual differences in music reward experiences. *Music Percept, 31*, 118-138 (2013).

Mattsson, B., Sorenson, J. C., Zimmer, J. & Johansson, B. B. Neural gifting to experimental neocortical infarcts improves behavioural outcome and reduces thalamic atrophy in rats housed in enriched but not in standard environments. *Stroke, 28*, 1225-1232 (1997).

Maylor, E. A. Serial position effects in semantic memory: Reconstructing the order of verses of hymns. *Psychon Bull Rev, 9*, 816-820 (2002).

McDonald, C. R., Ahmadi, M. E., Hagler, D. J., Tecoma, E. S., Iragui, V. J., Gharapetian, L., Dale, A. M. & Halgren, E. Diffusion tensor imaging correlates of memory and language impairments in temporal lobe epilepsy. *Neurology, 71,* 1869-1876 (2008).

McElhinney, M., & Annett, J. M. Pattern of efficacy of a musical mnemonic on recall of familiar words over several presentations. *Percept Mot Ski, 82*, 395-400 (1996).

Méndez Orellana, C. P., van de Sandt-Koenderman, M. E., Saliasi, E., van der Meulen, I., Klip, S., van der Lugt, A. & Smits, M. Insight into the neurophysiological processes of melodically intoned language with functional MRI. *Brain Behav, 4*, 615-25 (2014).

Merrett, D.L., Peretz, I. & Wilson, S. J. Neurobiological, cognitive, and emotional mechanisms in melodic intonation therapy. *Front Hum Neurosci, 8*, 401 (2014).

Moore, S. A., Hallsworth, K., Jakovljevic, D. G., Blamire, A. M., He, J., Ford, G. A., Rochester, L. & Trenell, M. I. Effects of community exercise therapy on metabolic, brain, physical, and cognitive function following stroke: a randomized controlled pilot trial. *Neurorehabil Neural Repair*, *29*, 623-35 (2015).

Morris, J. C., Heyman, A., Mohs, R. C., Hughes, J. P., van Belle, G., Fillenbaum, G., Mellits, E. D. & Clark, C. The consortium to establish a registry for alzheimer's disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology*, *39*, 1159-65 (1989).

Moussard, A., Bigand, E., Belleville, S. & Peretz, I. Learning sung lyrics aids retention in normal ageing and Alzheimer's disease. *Neuropsychol Rehabil*, *24,* 894-917 (2014).

Musso, M., Weiller, C., Kiebel, S., Müller, S. P., Bülau, P. & Rijntjes, M. Training-induced brain plasticity in aphasia. *Brain*, 122, 1781-90 (1999).

Neural basis of acquired amusia and its recovery. Sihvonen, A. J. Doctoral Thesis Turku: University of Turku (2018).

Neuropsychological Assessment. Lezak, M. D., Howieson, D. B. & Loring, D. W. Oxford: University Press (2004), 480-497; 518-20.

Nithianantharajah, J. & Hannan, A. J. Enriched environments, experience-dependent plasticity and disorders of the nervous system. *Nat Rev Neurosci, 7*, 697-709 (2006).

Norton, A., Zipse, L., Marchina, S. & Schlaug, G. Melodic intonation therapy: shared insights on how it is done and why it might help. *Ann N Y Acad Sci, 1169,* 431-6 (2009).

Nys, G. M., van Zandvoort, M. J., de Kort, P. L. & Jansen, B. P. Domain-specific cognitive recovery after first-ever stroke: A follow-up study of 111 cases. *J Int Neuropsychol Soc, 11*, 795-806 (2005).

Nys, G. M., Van Zandvoort, M. J., De Kort, P. L., Jansen, B. P., de Haan, E. H. & Kappelle, L. J. Cognitive disorders in acute stroke: prevalence and clinical determinants. *Cerebrovasc Dis., 23*, 408-16 (2007).

Ostergaard, A. L., & Meudell, P. R. Immediate memory span, recognition memory for subspan series of words, and serial position effects in recognition memory for supraspan series of verbal and nonverbal items in Broca's and Wernicke's aphasia. *Brain Lang, 22*, 1-13 (1984).

Palisson, J., Roussel-Baclet, C., Maillet, D., Belin, C., Ankri, J., & Narme, P. Music enhances verbal episodic memory in Alzheimers disease. *J Clin Exp Neuropsychol, 37*, 503-517 (2015).

Park, J. Y., Kim, Y. H., Chang, W. H., Park, C. H., Shin, Y. I., Kim, S. T. & Pascual-Leone, A. Significance of longitudinal changes in the default-mode network for cognitive recovery after stroke. *Eur J Neurosci*, 40, 2715-22 (2014).

Patel, A. D. Language, music, syntax and the brain. *Nat Neurosci*, *6*, 674-81 (2003).

Paulesu, E., Frith, C. D. & Frackowiak, R. S. The neural correlates of the verbal component of working memory. *Nature, 362*, 42–345 (1993).

Pereira, C. S., Teixeira, J., Figueiredo, P., Xavier, J., Castro, S. L. & Brattico, E. Music and emotions in the brain: familiarity matters. *PLoS One, 6*, e27241 (2011).

Peretz, I., Champod, A. S. & Hyde, K. Varieties of musical disorders. The Montreal Battery of Evaluation of Amusia. *Ann N Y Acad Sci, 999*, 58-75 (2003).

Peretz, I., & Coltheart, M. Modularity of music processing. *Nat Neurosci, 6*, 688-691 (2003).

Peretz, I., Gagnon, L., Hébert, S., & Macoir, J. (2004). Singing in the brain: Insights from cognitive neuropsychology. *Music Percep, 21*, 373–390.

Pihlaja, R., Uimonen, J., Mustanoja, S., Tatlisumak, T. & Poutiainen, E. Post-stroke fatigue is associated with impaired processing speed and memory functions in first-ever stroke patients. *J Psychosom Res, 77*, 380-384 (2014).

Platel, H., Baron, J. C., Desgranges, B., Bernard, F. & Eustache F. Semantic and episodic memory of music are subserved by distinct neural networks. *Neuroimage, 20*, 244-56 (2003).

Pongan, E, Tillmann, B., Leveque, Y., Trombert, B., Getenet, J. C., Auguste N., Dauhinot, V., El Haouari, H., Navez, M., Dorey, J. M., Krolak-Salmon, P., The LACMé Group, Laurent, B. & Rouch, I. Can musical or painting interventions improve chronic pain, mood, quality of life, and cognition in patients with mild Alzheimer's disease? Evidence from a randomized control trial. *J. Alzheimer's Dis.* 60, 663–677 (2017).

Progress in brain research. Finn, S. & Fancourt, D. In Christensen, J. F., Gomila, A. (Eds.). Amsterdam: Elsevier, 2018, 173–200.

Purnell-Webb, P. & Speelman, C. P. Effects of music on memory for text. *Percept Mot Ski, 106*, 927–957, (2008).

Quinlan, C. K. & Taylor, T. L. Enhancing the production effect in memory. *Memory, 21*, 904-915 (2013).

Racette, A., Bard, C., & Peretz I. Making non-fluent aphasics speak: sing along! *Brain, 129*, 2571-2784 (2006).

Racette, A., & Peretz, I. Learning lyrics: to sing or not to sing? *Mem Cognit, 35*, 242-253 (2007).

Raglio, A., Oasi, O., Gianotti, M., Rossi, A., Goulene, K. & Stramba-Badiale, M. Improvement of spontaneous language in stroke patients with chronic aphasia treated with music therapy: a randomized controlled trial. *International J Neurosci, 126*, 235-42 (2016).

Raichle, M. E. The brain's default mode network. *Annu Rev Neurosci, 38*, 433–447 (2015).

Rainey, D. W., & Larsen, J. D. The effect of familiar melodies on initial learning and long-term memory for unconnected text. *Music Percept, 20*, 173-186 (2002).

Ramsey, L. E., Siegel, J. S. Lang, C. E., Strube, M., Shulman, G. L. & Corbetta, M. Behavioural clusters and predictors of performance during recovery from stroke. *Nat Hum Behav, 1*, 0038 (2017).

Rathore, S. S., Hinn, A. R., Cooper, L. S., Tyroler, H. A., & Rosamond, W. D. Characterization of incident stroke signs and symptoms findings from the atherosclerosis risk in communities study. *Stroke*, *33*, 2718–2721 (2002).

Recovery after stroke. Diserens & Rothacher in M. P. Barnes, M. P., Dobkin B. H. & Bogousslavsky, J. (Eds.). Cambridge: Cambridge University Press (2005), 189-208.

Reggente, N., Cohen, M. S., Zheng, Z. S., Castel, A. D., Knowlton, B. J., & Rissman, J. Memory recall for high reward value items correlates with individual differences in white matter pathways associated with reward processing and fronto-temporal communication. *Front Hum Neurosci, 12*, 241 (2018).

Rentfrow, P. J., Goldberg, L. R. & Levitin, D. J. The Structure of Musical Preferences: A Five-Factor Model. *J Pers Soc Psychol, 100*, 1139–1157 (2011).

Revonsuo, A. & Portin, R. *CogniSpeed*: the computer-based measurement of cognitive processing. Turku, Finland: AboaTech; (1995).

Ripollés, P., Ferreri, L., Mas-Herrero, E., Alicart, H., Gómez-Andrés, A., Marco-Pallares, J., Antonijoan, R. M., Noesselt, T., Valle, M., Riba, J. & Rodriguez-Fornells, A. Intrinsically regulated learning is modulated by synaptic dopamine signaling. *Elife*, *7*, e38113 (2018).

Ripollés, P., Marco-Pallares, J., de Diego-Balaguer, R., Miro, J., Falip, M., Juncadella, M., Rubio, F. & Rodriguez-Fornells, A. Analysis of automated methods for spatial normalization of lesioned brains. *Neuroimage*, *60*, 1296-1306 (2012).

Ripollés, P., Rojo, N., Grau-Sánchez, J., Amengual, J., Càmara, E., Marco-Pallarés, J., Juncadella, M., Vaquero, L., Rubio. F., Duarte. E., Garrido, C., Altenmüller, E., Münte, T. Rodríguez-Fornells, A. Music supported therapy promotes motor plasticity in individuals with chronic stroke. *Brain Imaging Behav, 10*, 1289-307 (2016).

Rodd, J. M., Vitello, S., Woollams, A. M. & Adank, P. Localising semantic and syntactic processing in spoken and written language comprehension: an activation likelihood estimation meta-analysis. *Brain Lang, 141*, 89-102 (2015).

Rogenmoser, L., Zollinger, N., Elmer, S. & Jäncke, L. Independent component processes underlying emotions during natural music listening. *Soc Cogn Affect Neurosci*, 11, 1428-39 (2016).

Rorden, C. & Brett, M. Stereotaxic display of brain lesions. *Behav Neurol,* 12, 191-200 (2000).

Rosbergen, I. C., Grimley, R. S., Hayward, K. S., Walker, K. C., Rowley, D., Campbell, A. M., McGufficke, S., Robertson, S. T., Trinder, J., Janssen, H., Brauer, S. Embedding an enriched environment in an acute stroke unit increases activity in people with stroke: a controlled before-after pilot study. *Clin Rehabil, 31*, 1516-28 (2017).

Rosenthal, M. A. Hemispheric asymmetry in the hierarchical perception of music and speech. *Psychol Bull, 142*, 1165-78 (2016).

Royal, I., Vuvan, D. T., Zendel, B. R., Robitaille, N., Schönwiesner, M. & Peretz, I. Activation in the right inferior parietal lobule reflects the representation of musical structure beyond simple pitch discrimination. *PLoS One, 11,* e0155291 (2016).

Rubin, D. C. Very long-term memory for prose and verse. *J Verbal Learning Verbal Behav, 16*, 611-621 (1977).

Salamé, P. & Baddeley, A. Effects of background music on phonological short-term memory. *Q J Exp Psychol [A], 41*, 107-122 (1989).

Salimpoor, V. N., van der Bosch, I., Kovacevic, N., McIntosh, A. R., Dagher, A. & Zatorre, R. J. Interactions between the nucleus accumbens and auditory cortices predict music reward value. *Science, 340*, 216-219 (2013).

Salmi, J., Koistinen, O.-P., Glerean, E., Jylänki, P., Vehtari, A., Jääskeläinen, I. P., Mäkelä, S., Nummenmaa, L., Nummi-Kuisma, K., Nummi, I. & Sams, M. Distributed neural signatures of natural audiovisual speech and music in the human auditory cortex. Neuroimage, 157, 108-117. (2017).

Sammler, D., Baird, A., Valabregue, R., Clement, S., Dupont, S., Belin, P. & Samson, S. The relationship of lyrics and tunes in the processing of unfamiliar songs: a functional magnetic resonance adaptation study. *J Neurosci, 30*, 3572-8 (2010).

Samson, S., & Zatorre, R. J. Recognition memory for text and melody for songs after unilateral temporal lobe lesion: Evidence for dual encoding. *J Exp Psychol Learn Mem Cogn, 17,* 793-804 (1991).

Sarubbo, S., De Benedictis, A., Maldonado, I. L., Basso, G. & Duffau, H. Frontal terminations for the inferior fronto-occipital fascicle: anatomical dissection, DTI study and functional considerations on a multi-component bundle. *Brain Struct Funct, 218*, 21-37 (2013).

Saur, D., Lange, R., Baumgaertner, A., Schraknepper, V., Willmes, K., Rijntjes, M. & Weiller, C. Dynamics of language reorganization after stroke. *Brain*, 129, 1371-84 (2006).

Schauer, M. & Mauritz, K. Musical motor feedback (MMF) in walking hemiparetic stroke patients: randomized trials of gait improvement. *Clin Rehabil, 17*, 713-22 (2003).

Schlaug, G., Marchina, S., & Norton, A. From singing to speaking: Why singing may lead to recovery of expressive language function in patients with Broca's aphasia. *Music Percept, 25,* 315-323 (2008).

Schlaug, G., Marchina, S., & Norton, A. Evidence for plasticity in white-matter tracts of patients with chronic Broca's aphasia undergoing intense intonation-based speech therapy. *Ann N Y Acad Sci, 1169*, 385-394 (2009).

Schneider, S., Münte, T., Rodriguez-Fornells, A., Sailer, M. & Altenmüller, E. Music-supported training is more efficient than functional motor training for recovery of fine motor skills in stroke patients. *Music Percept, 27*, 271-80 (2010).

Schuppert, M., Thomas F. Münte, T. F., Wieringa, B. M. & Altenmüller, E. Receptive amusia: evidence for cross-hemispheric neural networks underlying music processing strategies. *Brain, 123*, 546–559 (2000).

Schön, D., Gordon, R., Campagne, A., Magne, C., Astésano, C., Anton, J. L. & Besson, M. Similar cerebral networks in language, music and song perception. *Neuroimage*, *51*, 450-61 (2010).

Schön, D., Boyer, M., Moreno, S., Besson, M., Peretz, I., & Kolinsky, R. Songs as an aid for language acquisition. *Cognition, 106*, 975-983 (2008).

Segado, M., Hollinger, A., Thibodeau, J., Penhune, V. & Zatorre, R. J. Partially overlapping brain networks for singing and cello playing. *Front Neurosci, 12*, 351 (2018).

Sihvonen, A. J., Särkämö, T., Leo, V., Tervaniemi, M., Altenmüller, E. & Soinila, S. Music-based interventions in neurological rehabilitation. *Lancet Neurol*, 16, 648-60. (2017a).

Sihvonen, A. J., Ripollés, P., Leo, V., Rodríguez-Fornells, A., Soinila, S. & Särkämö, T. Neural basis of acquired amusia and its recovery after stroke. *J Neurosci, 36*, 8872-81 (2016).

Sihvonen, A. J., Särkämö, T., Ripollés, P., Leo, V., Saunavaara, J., Parkkola, R., Rodríguez-Fornells, A. & Soinila, S. Functional neural changes associated with acquired amusia across different stages of recovery after stroke. *Sci Rep, 7*, 11390 (2017b).

Sihvonen, A. J., Ripolles, P., Särkämö, T., Leo, V., Rodriguez-Fornells, A., Saunavaara, J., Parkkola, R., Soinila, S. Tracting the neural basis of music:

Deficient structural connectivity underlying acquired amusia. Cortex, 97, 255-273 (2017c).

Sikka, R., Cuddy, L. L., Johnsrude, I. S. & Vanstone, A. D. An fMRI comparison of neural activity associated with recognition of familiar melodies in younger and older adults. *Front Neurosci*, 9, 356 (2015).

Silverman, M. J. The effect of paired pitch, rhythm, and speech on working memory as measured by sequential digit recall. *J Music Ther, 44*, 415-427 (2007).

Simmons-Stern, N. R., Budson, A. E., & Ally, B. A. Music as a memory enhancer in patients with Alzheimer's disease. *Neuropsychologia, 48*, 3164-3167 (2010).

Smith, S. M. Fast robust automated brain extraction. *Hum Brain Mapp, 17*, 143-155 (2002).

Smith, S. M., Fox, P. T., Miller, K. L., Glahn, D. C., Fox, P. M., Mackay, C. E., Filippini, N.., Watkins, K. E., Toro, R., Laird, A. R. & Beckmann, C. F. Correspondence of the brain's functional architecture during activation and rest. *Proc Natl Acad Sci U S A, 106*, 13040-5 (2009).

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E., Johansen-Berg, H., Bannister, P. R., De Luca, M., Drobnjak, I., Flitney, D. E., Niazy, R. K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J. M. & Matthews, P. M. Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage, 23*, 208-19 (2004).

Spalletta, G., Cravello, L., Gianni, W., Piras, F., Iorio, M., Cacciari, C., Casini, A. R., Chiapponi, C., Sancesario, G., Fratangeli, C., Orfei, M. D., Caltagirone, C. & Piras, F. Homotaurine effects on hippocampal volume loss and episodic memory in amnestic mild cognitive impairment. *J Alzheimers Dis, 50*, 807-816 (2016).

Staffaroni, A. M., Brown, J. A., Casaletto, K. B., Elahi, F. M., Deng, J., Neuhaus, J., Cobigo, Y., Mumford, P. S., Walters, S., Saloner, R., Karydas, A., Coppola, G., Rosen, H. J., Miller, B. L., Seeley, W. W. & Kramer, J. H. The longitudinal trajectory of default mode network connectivity in healthy older adults varies as a function of age and as associated with changes in episodic memory and processing speed. *J. Neurosci, 38*, 2809–2817 (2018).

Staffaroni, A. M., Melrose, R. J., Leskin, L. P., Riskin-Jones, H., Harwood, D., Mandelkern, M. & Sultzer, D. L. The functional neuroanatomy of verbal memory in Alzheimer's disease: [18F]-Fluoro-2-deoxy-D-glucose positron emission tomography (FDG-PET) correlates of recency and recognition memory. *J Clin Exp Neuropsychol, 39*, 682-693 (2017).

Stebbins, G. T., Nyenhuis, D. Ll, Wang, C., Cox, J. L., Freels, S., Bangen, K., deToledo-Morrell, L., Sripathirathan, K., Moseley, M., Turner, D. A., Gabrieli, J. D. &

Gorelick, P. B. Gray matter atrophy in patients with ischemic stroke with cognitive impairment. *Stroke* 39: 785– 793 (2008).

Stewart, L., von Kriegstein, K., Warren, J. D. & Griffiths, T. D. Music and the brain: Disorders of musical listening. *Brain, 129*, 2533–2553 (2006).

Straube, T., Schulz, A., Geipel, K., Mentzel, H.-J., & Miltner, W. H. R. Dissociation between singing and speaking in expressive aphasia: The role of song familiarity. *Neuropsychologia, 46*, 1505-1512 (2008).

Stroke. Current Care Guidelines. Working group set up by the Finnish medical society Duodecim and the Finnish cardiac society. Helsinki: The Finnish Medical Society Duodecim, retrieved November 24, 2019 from www.kaypahoito.fi (2016).

Su, M., Zhao, J., de Thiebaut S., Zhou, W., Gong, G., Ramus, F & Shu, H. Alterations in white matter pathways underlying phonological and morphological processing in Chinese developmental dyslexia. *Dev Cogn Neurosci, 31,* 11-19 (2018).

Särkämö, T., Ripollés, P., Vepsäläinen, H. R., Autti, T. H., Silvennoinen, H. M., Salli, E., Laitinen, S., Forsblom, A., Soinila, S. & Rodríguez-Fornells, A. Structural changes induced by daily music listening in the recovering brain after middle cerebral artery stroke: a voxel-based morphometry study. *Front Hum Neurosci, 8*, 245 (2014).

Särkämö, T., Pihko, E., Laitinen, S., Forsblom, A., Soinila, S., Mikkonen, M., Autti, T., Silvennoinen, H. M., Erkkilä, J., Lainen, M., Peretz, I., Hietanen, M. & Tervaniemi, M. Music and speech listening enhance the recovery of early sensory processing after stroke. *J Cogn Neurosci, 22*, 2716-27 (2010).

Särkämö, T & Sihvonen, A. J. Golden oldies and silver brains: Deficits, preservation, learning, and rehabilitation effects of music in ageing-related neurological disorders. *Cortex, 109*, 104-123 (2018).

Särkämö, T. & Soto, D. Music listening after stroke: beneficial effects and potential neural mechanisms. *Ann N Y Acad Sci, 1252*, 266-81 (2012).

Särkämö, T., Tervaniemi, M. & Huotilainen, M. Music perception and cognition: development, neural basis, and rehabilitative use of music. *Wiley Interdiscip Rev Cogn Sci, 4*, 441-451 (2013).

Särkämö, T., Tervaniemi, M., Laitinen, S., Forsblom, A., Soinila, S., Mikkonen, M., Autti, T., Silvennoinen, H. M., Erkkilä, J., Laine, M., Peretz I. & Hietanen, M. Music listening enhances cognitive recovery and mood after middle cerebral artery stroke. *Brain*, *131*, 866-76 (2008).

Särkämö, T., Tervaniemi, M., Soinila, S., Autti, T., Silvennoinen, H., Laine, M. & Hietanen, M. Cognitive deficits associated with acquired amusia after stroke: A neuropsychological follow-up study. *Neuropsychologia, 47*, 2642–2651 (2009).

Tabei, K. Inferior frontal gyrus activation underlies the perception of emotions, while precuneus activation underlies the feeling of emotions during music listening. *Behav Neurol, 2015*, 529043 (2015).

Tamminen, J., Rastle, K., Darby, J., Lucas, R., & Williamson, V. J. The impact of music in learning and consolidation of novel words. *Memory, 25*, 107-121 (2017).

Tang Q, Li G, Liu T, Wang A, Feng S, Liao, X., Jin, Y., Guo, Z., He, B., McClure, M. A., Xing, G. & Mu, Q. Modulation of interhemispheric activation balance in motor-related areas of stroke patients with motor recovery: Systematic review and meta-analysis of fMRI studies. *Neurosci Biobehav Rev*, *57*, 392-400 (2015).

Taruffi, L., Pehrs, C., Skouras, S. & Koelsch, S. Effects of sad and happy music on mind-wandering and the default mode network. *Sci Rep*, 7, 14396 (2017).

Tervaniemi, M. & Hugdahl, K. Lateralization of auditory-cortex functions. *Brain Res Rev.*, *43*, 231-46 (2003).

Thaut, M. H., Leins, A. K., Rice, R. R., Argstatter, H., Kenyon, G. P., McIntosh, G. C., Bolay, H. V. & Fetter, M. Rhythmic auditory stimulation improves gait more than NDT/Bobath training in near-ambulatory patients early poststroke: a single-blind, randomized trial. *Neurorehabil Neural Repair, 21*, 455-9 (2007).

Thaut, M. H., Peterson, D. A., & McIntosh, G. C. Temporal entrainment of cognitive functions: Musical mnemonics induce brain plasticity and oscillatory synchrony in neural networks underlying memory. *Ann N Y Acad Sci, 1060*, 243-254 (2005).

Thaut, M. H., Peterson, D.A., McIntosh, G. C., & Hoemberg, V. Music mnemonics aid verbal memory and induce learning-related brain plasticity in multiple sclerosis. *Front Hum Neurosci, 8*, 395 (2014).

The Assessment of aphasia and related disorders, 2nd ed. Goodglass, H. & Kaplan, E. Philadelphia: Lea & Febiger (1983).

The Oxford Handbook of Memory. Crowder, R. G. & Greene, R. L. in Tulving, E. & Craik, F. I. (Eds.). Oxford, England: Oxford University Press, 2000, 125-135.

The psychology of learning and motivation: Advances in research and theory. K. W. Spence & J. T. Spence (eds.) New York: Academic Press, 1968, 89–195.

The Rivermead Behavioral Memory Test. Wilson, B. A., Cockburn, J., & Baddeley, A. Bury St-Edmunds, England: Thames Valley Company (1985).

Thiebaut de Schotten, M., Cohen, L., Amemiya, E., Braga, L. W. & Dehane, S. Learning to read improves the structure of the arcuate fasciculus. *Cereb Cortex, 4*, 989-995 (2014).

Tuladhar, A. M., Snaphaan, L., Shumskaya, E., Rijpkema, M., Fernandez, G., Norris, D. G. & de Leeuw, F. E. Default mode network connectivity in stroke patients. *PLoS One, 8*, e66556 (2013).

Turkeltaub, P. E. & Coslett, H. B. Localization of sublexical speech perception components. *Brain Lang, 114*, 1-15 (2010).

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer B. & Joliot, M. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*, *15*, 273-89 (2002).

Umarova, R. M., Nitschke, K., Kaller, C. P. , Klöppel, S., Beume, L., Mader, I., Martin, M., Hennig, J. & Weiller, C. Predictors and signatures of recovery from neglect in acute stroke. *Ann Neurol*, *79*, 673-86 (2016).

van den Bosch, I., Salimpoor, V. N., and Zatorre, R. J. Familiarity mediates the relationship between emotional arousal and pleasure during music listening. *Front Hum Neurosci, 7*534 (2013).

van der Meulen, I., van de Sandt-Koenderman, W. M., Heijenbrok-Kal, M. H., Visch-Brink, E. G. & Ribbers, G. M. The efficacy and timing of melodic intonation therapy in subacute aphasia. *Neurorehabil Neural Repair, 28*, 536-44 (2014).

van Vugt, F. T., Kafczyk, T., Kuhn, W., Rollnik, J. D., Tillmann, B. & Altenmüller, E. The role of auditory feedback in music-supported stroke rehabilitation: a single-blinded randomised controlled intervention. *Restor Neurol Neurosci, 34*, 297-311 (2016).

Wallace, W. T. Memory for music: Effect of melody on recall of text. *J Exp Psychol Learn Mem Cogn*, *20*, 1071-1085 (1994).

Wan, C. Y., Rüüber, T., Hohmann, A. & Schlaug, G. The therapeutic effects of singing in neurological disorders. *Music percept, 27*, 287-295 (2010).

Wan, C. Y., Zheng, X., Marchina, S., Norton, A. & Schlaug G. Intensive therapy induces contralateral white matter changes in chronic stroke patients with Broca's aphasia. *Brain Lang, 136*, 1-7 (2014).

Wang, R., Benner, T., Sorensen, A. G. & Wedeen, V. J. Diffusion toolkit: A software package for diffusion imaging data processing and tractography. *Proc Intl Soc Mag Reson Med, 15*, 3720 (2007).

Weiss, M. W., Trehub, S. E. & Schellenberg, E. G. Somethig int the way she sings: Enhanced memory for vocal melodies. *Psychol Sci, 23*, 1074-1078 (2012).

Whitehead, J. C. & Armony, J. L. Singing in the brain: neural representation of music and voice as revealed by fMRI. *Hum Brain Mapp 39*, 4913–4924 (2018).

Wilbiks, J. M., Vuvan, D. T., Girard, P.-Y., Peretz, I., & Russo, F. A. Effects of vocal training in a musicophile with congenital amusia. *Neurocase, 22,* 526–537 (2016).

Wilkins, R. W., Hodges, D. A., Laurienti, P. J., Steen, M. & Burdette, J. H. Network science and the effects of music preference on functional brain connectivity: from Beethoven to Eminem. *Sci Rep, 4*, 6130 (2014).

Yalch, R.F. Memory in a jingle jungle: music as a mnemonic device in communicating advertising slogans. *J Appl Soc Psychol, 76*, 268–275 (1991).

Yang, M., Yang, P., Fan, Y. S., Li, J., Yao, D., Liao, W. & Chen, H. Altered structure and intrinsic functional connectivity in post-stroke aphasia. *Brain Topogr*, *31*, 300-10 (2018).

Yang, Y., Chen, Y., Chen, K., Wei, D., Li, P., Zeng, W., Pei, J., Mao, H., Jia, J. & Zhang, Z. Increased intrinsic connectivity for structural atrophy and functional maintenance after acute ischaemic stroke. *Eur J Neurol, 26*, 935-942 (2019).

Zarate, J. M. The neural control of singing. *Front Hum Neurosci, 7*, 237 (2013).

Zatorre, R. J., Belin, P. & Penhune, V. B. Structure and function of auditory cortex: music and speech. *Trends Cogn Sci*, *6*, 37-46 (2002).

Zatorre, R. J., Chen, J. L. & Penhune, V. B. When the brain plays music: auditory-motor interactions in music perception and production. *Nat Rev Neurosci*, 8, 547-58 (2007).

Zatorre, R. J., Fields, R. D. & Johansen-Berg H. Plasticity in grey and white: neuroimaging changes in brain structure during learning. *Nat Neurosci, 15*, 528-536 (2012).

Zatorre, R. J. & Salimpoor, V.N. From perception to pleasure: music and its neural substrates. *Proc Natl Acad Sci U S A, 110*, S10430-7 (2013).

Zhang, J., Andreano, J. M., Dickerson, B. C., Touroutoglou, A. & Barrett, L. F. Stronger functional connectivity in the default mode and salience networks is associated with youthful memory in superaging. *Cereb Cortex*, pii: bhz071 (2019).

Zipse, L., Norton, A., Marchina, S. & Schlaug, G. When right is all that's left: Plasticity of the right hemisphere tracts in a young aphasic patient. *Ann N Y Acad Sci, 1252*, 237-245 (2012).

Zumbansen, A., Peretz, I. & Hébert, S. (2014a). Melodic intonation therapy: Back to basics for future research. *Front Neurol, 5*, 7.

Zumbansen, A., Peretz, I., & Hébert, S. (2014b). The combination of rhythm and pitch can account for the beneficial effect of melodic intonation therapy on connected speech improvements in Broca's aphasia. *Front Hum Neurosci, 8*, 592.

Özdemir, E., Norton, A., & Schlaug, G. Shared and distinct neural correlates of singing and speaking. *Neuroimage, 33*, 628-635 (2006).